Spectral Element Method with a Transparent Boundary Operator for Quasi-Periodic Helmholtz Solutions on Rough Structures

Ying He
Department of Mathematics
Purdue University
West Lafayette, IN 47907

Misun Min
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, IL 60439

David P. Nicholls
Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago
Chicago, IL 60607

February 9, 2014

Abstract

We present a transparent boundary operator for a high-order spectral element approach for solving exterior scattering problems governed by a scalar Helmholtz equation. In particular, we consider incident waves at arbitrary angles impinging on scattering surfaces with periodic gratings, where the scattering solutions are represented in a quasi-periodic form. We rewrite our governing equation into a formula that eliminates the quasi-periodicity and solve the reformulated scalar Helmholtz equation with periodic, Dirichlet, and transparent boundary conditions. We construct a spectral element Dirichlet-to-Neumann boundary operator for the transparent boundary condition that ensures nonreflecting outgoing waves on the artificial boundaries in the truncated computational domain. We present an explicit formula that accurately computes the Fourier data involved in the boundary operator on the spectral element discretization space. Our solutions are represented by the tensor product basis of the one-dimensional Legendre-Lagrange interpolation polynomials based on the Gauss-Lobatto-Legendre grids. We study scattered field solutions in single- and double-layer media with smooth and nonsmooth scattering surfaces. Geometric structures of the scattering surfaces include rectangular, triangular, and sawtooth grooves that are accurately represented by the body-fitted quadrilateral elements. We use a GMRES iteration technique to solve the resulting linear system. We validate our results provided with spectral convergence in comparison with exact solutions and the results by the transformed field expansion method, including the energy defect measure.

1 Introduction

Exterior scattering problems arise in many engineering applications in electromagnetics, optics, and acoustics. Robust and accurate simulation capability has been receiving an increased attention as a cost-effective tool for predictive measurement and analysis of modern physical systems. Highly accurate boundary treatment and flexibility to treat complex geometries are essential for solving exterior scattering problems arising in a broad range of applications.

Many competing numerical methods have been developed, such as the boundary integral and boundary element methods [1, 3]. These surface methods require discretization only at the interface, thus significantly reducing the number of unknowns to compute. Depending on the choice of the

1

Green's function, the far-field boundary conditions can be enforced exactly; and these methods can deliver highly accurate solutions with reduced operation counts. Such methods face a number of drawbacks, however, including the fact that the inhomogeneities away from the layer interfaces cannot be accommodated and high-order accuracy can be realized only with specially designed quadrature nodes, because of the singularities in the Green's function approach. Moreover, these methods typically give rise to a dense linear system that requires carefully designed preconditioned iterative methods with acceleration algorithms for the matrix-vector product, such as fast multipole methods [9].

As an alternative, boundary perturbation methods have been explored. Bruno & Reitich studied the method of field expansions [4, 5, 6], and Milder studied the method of operator expansions [11, 12, 13, 15, 16, 14]. These methods solve surface unknowns, thereby enjoying the favorable operation counts of surface integral methods, while avoiding the subtle quadrature rules, dense linear systems, and algorithms for matrix-vector product accelerations. However, these algorithms depend on strong cancellations, resulting in ill-conditioning [18, 19, 20]. Nicholls and Reitich proposed an enhanced boundary perturbation algorithm, referred to as the method of transformed field expansions (TFE) [10], which does not rely on strong cancellations. In this approach, the resulting recursions can be used for a direct, rigorous demonstration of the strong convergence of the relevant perturbation expansions in an appropriate function space. Furthermore, these formulas were proven to be a stable and accurate numerical scheme for simulating scattering problems defined on layered periodic gratings with a generalization to the case of irregularly bounded obstacles [21, 8], and a rigorous numerical analysis was provided in [22]. However, this method is limited when complex geometries are considered.

To address the limited simulation capabilities of the boundary methods, we consider a high-order spectral element method [7]. Specifically, in this paper we introduce a spectral element transparent boundary operator that has not been studied in the literature for exterior scattering problems. We focus on the two-dimensional formulation of the boundary operator in spectral element discretization space, utilizing the Dirichlet-to-Neumann (DtN) map approach [18, 19, 20], where the formulation involves the normal derivative of the truncated Fourier series of the Dirichlet data on the artificial boundaries located on the exterior domain [10]. We present an accurate formula for computing the Fourier data involved in the boundary operator in spectral element discretization space. In particular, we consider incident waves at arbitrary angles impinging on various types of surface periodic gratings, resulting in quasi-periodic solutions of the scalar Helmholtz equation. We rewrite our governing equation into a form that eliminates the quasi-periodicity and solve the reformulated scalar Helmholtz equation with periodic, Dirichlet, and transparent boundary conditions. We solve various example problems and demonstrate our computational results with validation. We note that the linear system featured with the quasi-periodicity of the problems and the spectral element DtN operator does not hold the positive definite Hermitian property. In this work, we simply use the generalized minimum residual (GMRES) method [24] without preconditioning for solving the linear system. We leave the study of efficient preconditioning techniques and algorithm extension to three dimensions as future efforts.

This paper is organized as follows. In Section 2, we define the governing equations for our model problems and provide formulations. Section 3 discusses the spectral element discretization. Section 4 presents the computational results and their validation. Section 5 summarizes our conclusion.

2   Formulations

A time-harmonic plane wave of frequency w can be expressed typically in a complex form as

$$\bar{U}(x,y,t) = U(x,y)e^{-\mathbf{i}wt} = e^{\mathbf{i}\kappa \cdot \mathbf{x}}e^{-\mathbf{i}wt} = e^{\mathbf{i}(\alpha x + \beta y)}e^{-\mathbf{i}wt}, \tag{2.1}$$

where the wave vector $\kappa = (\alpha, \beta)$ indicates the propagation direction in a homogeneous medium. The monochromatic waves $U(x,y)$ with different w satisfy a scalar Helmholtz equation given as

$$\Delta U + k^2 U = 0, \quad k = |\kappa| = \frac{w}{c}, \tag{2.2}$$

where the constant c is the wavespeed of the medium and the wavenumber k is the number of waves per 2π for real w. Note that both k and w can be either negative or positive, although the Helmholtz equation is invariant regardless of the sign change of k. Polychromatic waves can be the sums of waves of different frequencies by taking the forward Fourier transform of U, which relates to the inverse Fourier transform of $\bar{U}$ as follows:

$$\bar{U}(x,y,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\mathbf{i}wt}U(x,y)dw \quad \text{and} \quad U(x,y) = \int_{-\infty}^{\infty} e^{\mathbf{i}wt}\bar{U}(x,y,t)dt. \tag{2.3}$$

Here $\bar{U}$ is the solution of the wave equation, and thus the Helmholtz equation (2.2) can be considered as the frequency-domain formulation of the wave equation:

$$\frac{\partial^2 \bar{U}}{\partial t^2} - c^2 \Delta \bar{U} = 0. \tag{2.4}$$

In this paper, we consider the scalar Helmholtz equation for time-harmonic scattering problems.

2.1   Model Problems

We describe two classes of model problems. Consider exterior scattering problems in single- and double-layer media as shown in Figure 1. We define the unbounded domains,

$$\Omega_0^+ = \{y > g(x)\} \quad \text{and} \quad \Omega_0^- = \{y < g(x)\}, \tag{2.5}$$

with different wave numbers, k^+ and k^-, given in Ω_0^+ and Ω_0^-, respectively. We consider the scatterer boundary denoted by $\Gamma_g = \{(x,y) \in \mathbb{R}^2, \ y = g(x) \in L^2(\mathbb{R})\}$, which represents the shape of grating structure (smooth or rough) with d-periodicity, not necessarily a unit cell. Here we note that the total field solutions are quasi-periodic [2], in the sense that

$$U(x+d, y) = e^{\mathbf{i}\alpha d}U(x,y). \tag{2.6}$$

For the single-layer model shown in Figure 1(a), a homogeneous Dirichlet boundary is considered on Γ_g, using the notation $\Gamma_D := \Gamma_g$ for the Dirichlet boundary, and the scattered waves must be outgoing as $y \to +\infty$. For the double-layer model shown in Figure 1(b), the total field is required to be continuous across the scatterer interface Γ_g, and the scattered waves must be outgoing as $y \to \pm\infty$. These model problems can be described by the Helmholtz equation with proper boundary conditions defined as follows.

Model 1. The total field U in the single layer, denoting $\Omega_0 = \Omega_0^+$, is described as

$$\Delta U + k^2 U = 0 \qquad\qquad\qquad \text{on } \Omega_0, \tag{2.7}$$
$$U(x+d, y) = e^{\mathbf{i}\alpha d}U(x,y) \qquad\qquad \text{on } \Omega_0, \tag{2.8}$$
$$U(x,y) = 0 \qquad\qquad\qquad\qquad \text{on } \Gamma_D. \tag{2.9}$$

3

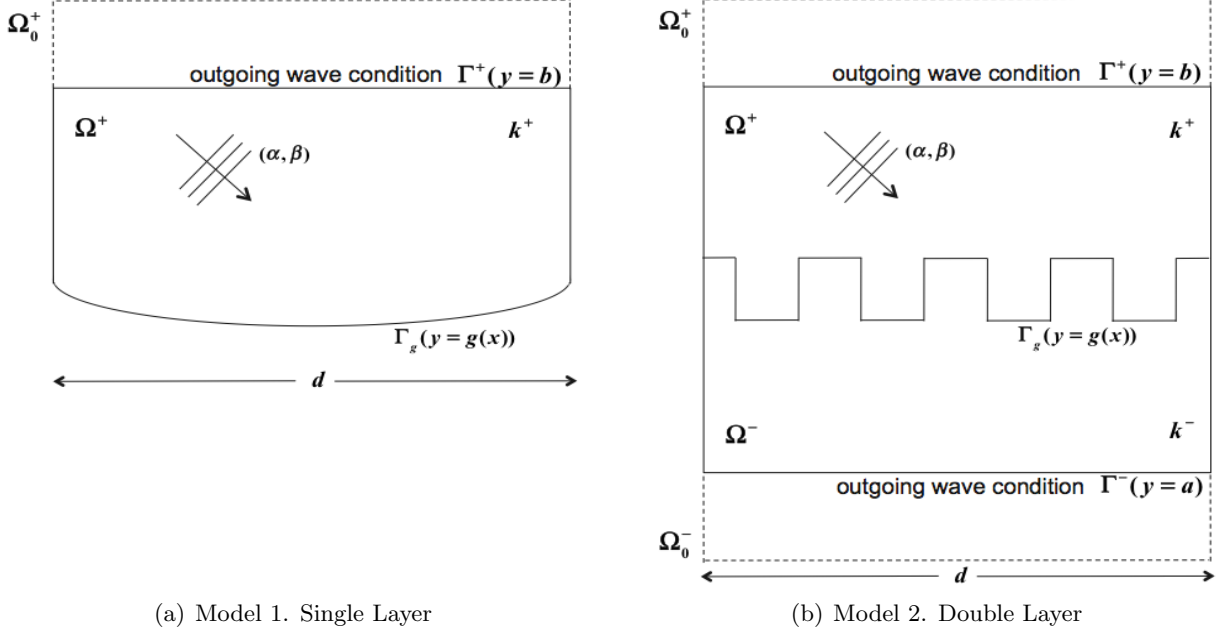(a) Model 1. Single Layer   (b) Model 2. Double Layer

Figure 1: Geometric illustration of the model problems.

Model 2. The total field U in the double layer, denoting $\Omega_0 = \Omega_0^+ \cup \Omega_0^- \cup \Gamma_g$, is described as

$$\Delta U + k^2 U = 0 \qquad\qquad\qquad\qquad \text{on } \Omega_0, \qquad (2.10)$$

$$U(x + d, y) = e^{\mathbf{i}\alpha d} U(x, y) \qquad\qquad \text{on } \Omega_0. \qquad (2.11)$$

2.2 Quasi-Periodic to Periodic Formulation

We consider incident waves at arbitrary angles with $\alpha \in \mathbb{R}$. For the normal incident wave ($\alpha = 0$), the solution U is still periodic. For an oblique incident wave ($\alpha \neq 0$), however, the solution is characterized as quasi-periodic in the x-direction. In our algorithm implementation, it is more convenient to handle the periodic boundary than the quasi-periodic boundary condition. Thus we introduce a new variable u as

$$u(x, y) = e^{-\mathbf{i}\alpha x} U(x, y), \qquad (2.12)$$

where u is periodic for any α from the fact that

$$u(x + d, y) = e^{-\mathbf{i}\alpha(x+d)} U(x + d, y) = e^{-\mathbf{i}\alpha(x+d)} e^{\mathbf{i}\alpha d} U(x, y) = u(x, y). \qquad (2.13)$$

Plugging (2.12) into Eqs. (2.7) and (2.10), we solve the reformulated scalar Helmholtz equation:

$$\Delta u + (k^2 - \alpha^2)u + 2\mathbf{i}\alpha \frac{\partial u}{\partial x} = 0. \qquad (2.14)$$

We note that the first-order derivative term in Eq. (2.14) results from the quasi-periodicity of the solution U with $\alpha \neq 0$; hence, we refer to it as a quasi-periodic term. We also note that in general $k^2 - \alpha^2$ can be either positive or negative, although we consider only positive case in this paper. These properties bring a different nature to our model problems from that of the problems involving the usual Helmholtz operator.

4

2.3 Transparent Boundary Conditions

From the separation of variables, we can express the general solution U for the Helmholtz equations (2.7) and (2.10) in a Fourier series. Then, applying the d-periodicity in x-direction through the relation (2.12), the solution u for the governing equation (2.14) on Ω_0^+ or Ω_0^- can be written as

$$u(x,y) = \sum_{p=-\infty}^{\infty} \hat{u}_p e^{\mathbf{i}(\alpha_p - \alpha)x} e^{\mathbf{i}\beta_p y}, \tag{2.15}$$

where $\hat{u}_p$ are the Fourier coefficients of u with $\alpha_p = \alpha + \frac{2\pi p}{d}$ and $\beta_p^2 = k^2 - \alpha_p^2$ for an integer p. Defining $\mathbb{K} = \left\{ p \in \mathbb{Z} \mid k^2 - \alpha_p^2 > 0 \right\}$, we can have

$$\beta_p = \pm\sqrt{k^2 - \alpha_p^2} \quad p \in \mathbb{K} \quad \text{and} \quad \beta_p = \pm\mathbf{i}\sqrt{\alpha_p^2 - k^2} \quad p \notin \mathbb{K}. \tag{2.16}$$

Assume an incident field $u_{\text{inc}} = e^{\mathbf{i}(\alpha x + \beta y)}$ defined on Ω_0^+. The total field solution u can be expressed as the summation of the incident and scattered fields as

$$u = u_{\text{inc}} + u_{\text{scat}}, \tag{2.17}$$

and thus the scattered field solution can be also in the form of (2.15) as

$$u_{\text{scat}} = \sum_{p=-\infty}^{\infty} \hat{u}_p^{\text{scat}} e^{\mathbf{i}(\alpha_p - \alpha)x} e^{\mathbf{i}\beta_p y}. \tag{2.18}$$

Consider the propagation direction $\kappa = (\alpha, \beta)$ with $\alpha > 0$ and $\beta < 0$ for an incident. Then, the outgoing waves are allowed only to have either propagating or evanescent modes with

$$\beta_p = \begin{cases} \sqrt{k^2 - \alpha_p^2} & p \in \mathbb{K} \\ \mathbf{i}\sqrt{\alpha_p^2 - k^2} & p \notin \mathbb{K} \end{cases} \quad \text{on} \ \ \Omega_0^+ \quad \text{and} \quad \beta_q = \begin{cases} -\sqrt{k^2 - \alpha_q^2} & q \in \mathbb{K} \\ -\mathbf{i}\sqrt{\alpha_q^2 - k^2} & q \notin \mathbb{K} \end{cases} \quad \text{on} \ \ \Omega_0^-. \tag{2.19}$$

Now, we discuss a boundary operator that maps the solution at infinity in the infinite domain to the solution on the boundary of a finite domain. Consider the model problems defined on the finite computational domains, as shown in Figure 1, using the following notations:

$$\Omega^+ = \{g(x) < y < b\} \quad \text{and} \quad \Omega^- = \{a < y < g(x)\}. \tag{2.20}$$

Here we define a hyperplane $\Gamma = \{(x,y) \in \mathbb{R}^2, \ \ y = c^*\}$ such that $\Gamma \cap \Gamma_g = \emptyset$, where the constant c^* can represent either a or b for our model problems. Taking the normal derivative of u_{scat} on Γ, we define a transparent boundary operator T:

$$T[u_{\text{scat}}]\,|_{y=c^*} = \frac{\partial u_{\text{scat}}}{\partial \mathbf{n}}\,|_{y=c^*} = \mathbf{n} \cdot \nabla u_{\text{scat}}\,|_{y=c^*}. \tag{2.21}$$

The operator T maps Dirichlet data to Neumann data on Γ; hence, it is referred as a Dirichlet-to-Neumann (DtN) operator. For our model problems, the operator T can be expressed as

$$T[u_{\text{scat}}]\,|_{y=c^*} = n_y \frac{\partial u_{\text{scat}}}{\partial y}\,|_{y=c^*} = n_y \sum_{p=-\infty}^{\infty} (\mathbf{i}\beta_p)\hat{u}_p^{\text{scat}} e^{\mathbf{i}(\alpha_p - \alpha)x} e^{\mathbf{i}\beta_p c^*}, \tag{2.22}$$

where the outward unit normal vector is $\mathbf{n} = (n_x, n_y) = (0, \pm 1)$ with a proper set of indices for p depending on the domain, either Ω_0^+ or Ω_0^-. On the other hand, the following continuity conditions should hold for the solutions of (2.14) on Ω_0^+ and Ω_0^-:

$$u: \quad \text{continuous across } \Gamma, \tag{2.23}$$

$$\frac{\partial u}{\partial \mathbf{n}}: \quad \text{continuous across } \Gamma. \tag{2.24}$$

The transparent boundary operator (2.22) satisfies the continuity conditions across Γ. Thus we can define the outgoing scattered solution on Γ in the reduced finite domains Ω^+ and Ω^- as

$$T[u - u_{\text{inc}}] = \frac{\partial(u - u_{\text{inc}})}{\partial \mathbf{n}}, \tag{2.25}$$

which can be equivalently written as

$$\partial_{\mathbf{n}} u - T[u] = \partial_y(u_{\text{inc}}) - T[u_{\text{inc}}]. \tag{2.26}$$

2.4 Governing Equations

Defining $\Gamma^+ = \Gamma \subset \Omega_0^+$ (if $c^* = b$) and $\Gamma^- = \Gamma \subset \Omega_0^-$ (if $c^* = a$), we can summarize our governing equations for our model problems as follows:

Model 1. For the single layer case, denoting $\Omega = \Omega^+$, our governing equations are

$$\Delta u + (k^2 - \alpha^2)u + 2\mathbf{i}\alpha\frac{\partial u}{\partial x} = 0 \qquad \text{on } \Omega, \tag{2.27}$$

$$u(x + d, y) = u(x, y) \qquad \text{on } \Omega, \tag{2.28}$$

$$\partial_{\mathbf{n}} u - T[u] = \rho \qquad \text{on } \Gamma^+, \tag{2.29}$$

$$u = 0 \qquad \text{on } \Gamma_D, \tag{2.30}$$

where $\rho = \partial_y(u_{\text{inc}}) - T[u_{\text{inc}}]$.

Model 2. For the double layer case, denoting $\Omega = \Omega^+ \cup \Omega^- \cup \Gamma_g$, our governing equations are

$$\Delta u + (k^2 - \alpha^2)u + 2\mathbf{i}\alpha\frac{\partial u}{\partial x} = 0 \qquad \text{on } \Omega, \tag{2.31}$$

$$u(x + d, y) = u(x, y) \qquad \text{on } \Omega, \tag{2.32}$$

$$\partial_{\mathbf{n}} u - T[u] = \rho \qquad \text{on } \Gamma^+, \tag{2.33}$$

$$\partial_{\mathbf{n}} u - T[u] = 0 \qquad \text{on } \Gamma^-, \tag{2.34}$$

where $\rho = \partial_y(u_{\text{inc}}) - T[u_{\text{inc}}]$ and the wavenumbers are defined as $k = k^+$ on Ω^+ and $k = k^-$ on Ω^-. Note that we consider the incident wave u_{inc} only defined on Ω^+.

2.5 Variational Formulation

In this section, we derive the variational formulations of our governing equations for the model problems (2.27)–(2.30) and (2.31)–(2.34). Consider a test function $v \in H^1(\Omega)$. Multiplying v to (2.27) and (2.31) and integrating them over Ω, whose boundary is denoted by $\partial\Omega$, we have

$$\int_\Omega \nabla u \nabla \overline{v} d\Omega - \int_{\partial\Omega} \mathbf{n} \cdot \nabla u \overline{v} dS - \int_\Omega (k^2 - \alpha^2)u\overline{v}d\Omega - \int_\Omega 2\mathbf{i}\alpha\frac{\partial u}{\partial x}\overline{v}d\Omega = 0. \tag{2.35}$$

6

The surface integrations with the boundary conditions applied on the single layer are

$$\int_{\partial\Omega} \mathbf{n} \cdot \nabla u \overline{v} dS = \int_{\Gamma^+} T[u]\overline{v} d\Gamma - \int_{\Gamma^+} \rho \overline{v} d\Gamma - \int_{\Gamma_D} \mathbf{n} \cdot \nabla u \overline{v} d\Gamma, \tag{2.36}$$

and those for the double layer are

$$\int_{\partial\Omega} \mathbf{n} \cdot \nabla u \overline{v} dS = \int_{\Gamma^+} T[u]\overline{v} d\Gamma - \int_{\Gamma^+} \rho \overline{v} d\Gamma + \int_{\Gamma^-} T[u]\overline{v} d\Gamma. \tag{2.37}$$

We seek a solution $u \in H^1(\Omega)$ such that

$$a(u,v) = \langle \rho, v \rangle \qquad \text{for all } v \in H^1(\Omega), \tag{2.38}$$

where the sesquilinear form $(\cdot)$ for each model is defined as follows.

Model 1. From (2.35) and (2.36), we have

$$a(u,v) = \int_{\Omega} \left(\nabla u \cdot \nabla \overline{v} - (k^2 - \alpha^2) u \overline{v} - 2i\alpha \frac{\partial u}{\partial x}\overline{v} \right) d\Omega - \int_{\Gamma^+} T[u]\overline{v} d\Gamma. \tag{2.39}$$

Model 2. From (2.35) and (2.37), we have

$$a(u,v) = \int_{\Omega} \left(\nabla u \cdot \nabla \overline{v} - (k^2 - \alpha^2) u \overline{v} - 2i\alpha \frac{\partial u}{\partial x}\overline{v} \right) d\Omega - \int_{\Gamma^+} T[u]\overline{v} d\Gamma - \int_{\Gamma^-} T[u]\overline{v} d\Gamma. \tag{2.40}$$

The linear functional $\langle \cdot \rangle$ in (2.38) is defined for both models as follows:

$$\langle \rho, v \rangle = \int_{\Gamma^+} \rho \overline{v} d\Gamma. \tag{2.41}$$

In particular, we define the following notations for the volume integrations:

$$A(u,v) = \int_{\Omega} \nabla u \nabla \overline{v} d\Omega, \quad B(u,v) = \int_{\Omega} u \overline{v} d\Omega, \quad C(u,v) = \int_{\Omega} \frac{\partial u}{\partial x}\overline{v} d\Omega, \tag{2.42}$$

and for the surface integrations:

$$T(u,v) = \int_{\Gamma} T[u]\overline{v} d\Gamma, \quad F(\rho,v) = \int_{\Gamma^+} \rho \overline{v} d\Gamma, \tag{2.43}$$

where $\Gamma = \Gamma^+$ for the single-layer and $\Gamma = \Gamma^+ \cup \Gamma^-$ for the double-layer case. We note that

$$T(u,v) = n_y \sum_{p=-\infty}^{\infty} i\beta_p \hat{u}_p \int_{\Gamma} e^{id_p x}\overline{v} dx = n_y \sum_{p=-\infty}^{\infty} i\beta_p \hat{u}_p \int_{\Gamma} \overline{e^{-id_p x}v} dx = n_y \sum_{p=-\infty}^{\infty} i\beta_p \hat{u}_p \overline{\hat{v}_p}, \tag{2.44}$$

where $d_p = \frac{2\pi p}{d} = \alpha_p - \alpha$ in Eq. (2.22). On the other hand, we have

$$\overline{T(v,u)} = n_y \sum_{p=-\infty}^{\infty} \overline{i\beta_p \hat{v}_p \overline{\hat{u}_p}} = n_y \sum_{p=-\infty}^{\infty} (\overline{i\beta_p}) \hat{u}_p \overline{\hat{v}_p}. \tag{2.45}$$

The boundary operator T might provide special properties for the following cases: if β_p is real for all p, $T(u,v) = -\overline{T(v,u)}$; if β_p is imaginary for all p, $T(u,v) = \overline{T(v,u)}$. In general, however, $T(u,v) \neq \overline{T(v,u)}$.

7

(a) Local Numbering (Unassembled Representation)

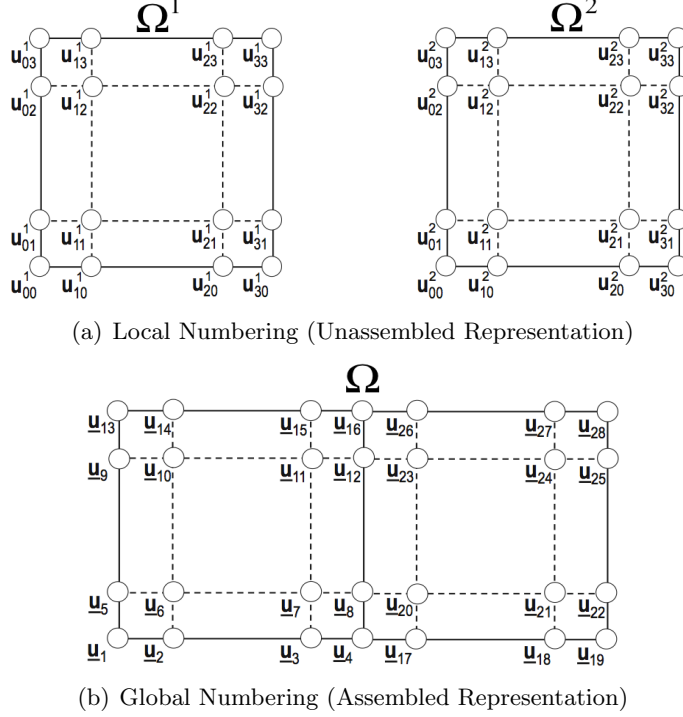(b) Global Numbering (Assembled Representation)

Figure 2: Illustration of a solution vector in a local numbering and a global numbering, using an example mesh with $(E, N) = (2, 3)$: $\Omega = \Omega^1 \cup \Omega^2$ and the GLL nodes ($\circ$).

3 Spectral Element Discretization

We denote our computational domain as $\Omega = \cup_{e=1}^{E} \Omega^e$, where Ω^e represents nonoverlapping body-conforming quadrilateral elements. Let us define a finite-dimensional approximation space $V_N \subset H^1(\Omega)$ such that $V_N = \text{span}\{\psi_{ij}(\xi, \eta)\}_{i,j=0}^{N}$. With this choice of approximation space, we consider a local approximate solution $u^e(x, y) \in V_N$, or simply u^e, that has the representation

$$u^e(x, y) = \sum_{i,j=0}^{N} u_{ij}^e \psi_{ij}(\xi, \eta). \tag{3.1}$$

The basis coefficients u_{ij}^e are the nodal values $u^e(x_i, y_j)$ on Ω^e, and the basis $\psi_{ij}(\xi, \eta) = \ell_i(\xi)\ell_j(\eta)$, or simply ψ_{ij}, has a tensor product form of the one-dimensional Nth-order Legendre-Lagrange interpolation polynomials given as

$$\ell_i(\xi) = [N(N+1)^{-1}(1-\xi^2)L_N'(\xi)]/[(\xi - \xi_i)L_N(\xi_i)] \quad \text{for } \xi \in [-1, 1], \tag{3.2}$$

based on the Gauss-Lobatto-Legendre (GLL) quadrature nodes ξ_i with the Nth-order Legendre polynomial L_N and its derivative L_N'. We map each physical coordinate $(x, y) \in \Omega^e$ onto the reference domain $(\xi, \eta) \in I = [-1, 1]^2$ through the Gordon-Hall mapping [7] and formulate computational scheme on the reference domain.

Let us denote our numerical solution $\mathbf{u}$ on Ω as with vector representations as

$$\mathbf{u} := (u_1, u_2, ..., u_{\hat{\imath}}, ..., u_n) := (u^1, u^2, ..., u^e, ..., u^E)^T, \tag{3.3}$$

$$u^e := (u_1^e, u_2^e, ..., u_l^e, ..., u_{(N+1)(N+1)}^e)^T := (u_{00}^e, u_{10}^e, ..., u_{ij}^e, ..., u_{NN}^e)^T, \tag{3.4}$$

8

where $n = E(N+1)^2$ is the total number of basis coefficients and $\hat{l} = 1+i+j(N+1)+(e-1)(N+1)^2$ and $l = 1+i+j(N+1)$ translate the two-index coefficient representation into a vector form, with the leading index i. In Figure 2, we show a mesh with two elements $E = 2$ including the GLL grids for $N = 3$ on $\Omega = \Omega^1 \cup \Omega^2$. Figure 2(a) illustrates a local ordering of a solution vector $\mathbf{u}$ based on the two-index expression in an unassembled representation for the coincident grids, $u^1_{3i} = u^2_{0i}$ ($i = 0, ..., 3$), appearing redundantly. In Figure 2(b), we demonstrate the same solution vector in a global ordering in an assembled representation using only the distinct nodes, denoted by

$$\underline{\mathbf{u}} = (\underline{u}_1, \underline{u}_2, ..., \underline{u}_{\bar{n}})^T. \tag{3.5}$$

The size ($\bar{n} < n$) of the solution vector $\underline{\mathbf{u}}$ in the assembled representation is reduced after eliminating the redundancy from the coincident grids. In practice, our implementations are based on elementwise computations using the data structure in the local ordering. The global ordering is used when it is more convenient to describe our method in this paper.

3.1 Stiffness Matrices

To obtain the stiffness matrix, we consider the following inner product in Eq. (2.42):

$$A(u, v) = \int_\Omega \nabla u \cdot \nabla \overline{v} d\Omega = \int_\Omega \left(\frac{\partial u}{\partial x} \frac{\partial \overline{v}}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial \overline{v}}{\partial y} \right) d\Omega, \tag{3.6}$$

where the partial derivatives are expressed by the chain rule for $x = x(\xi, \eta)$ and $y = y(\xi, \eta)$ on Ω^e:

$$
\begin{aligned}
\frac{\partial u}{\partial x} \frac{\partial \overline{v}}{\partial x} &= \left(\frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial x} \right) \left(\frac{\partial \overline{v}}{\partial \xi} \frac{\partial \xi}{\partial x} + \frac{\partial \overline{v}}{\partial \eta} \frac{\partial \eta}{\partial x} \right) \\
&= \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \xi} \left(\frac{\partial \xi}{\partial x} \frac{\partial \xi}{\partial x} \right) + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \eta} \left(\frac{\partial \eta}{\partial x} \frac{\partial \eta}{\partial x} \right) + \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \eta} \left(\frac{\partial \xi}{\partial x} \frac{\partial \eta}{\partial x} \right) + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \xi} \left(\frac{\partial \xi}{\partial x} \frac{\partial \eta}{\partial x} \right) \\
&= \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \xi} \mathcal{G}^{\xi\xi}_{xx} + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \eta} \mathcal{G}^{\eta\eta}_{xx} + \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \eta} \mathcal{G}^{\xi\eta}_{xx} + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \xi} \mathcal{G}^{\xi\eta}_{xx}, \tag{3.7} \\
\frac{\partial u}{\partial y} \frac{\partial \overline{v}}{\partial y} &= \left(\frac{\partial u}{\partial \xi} \frac{\partial \xi}{\partial y} + \frac{\partial u}{\partial \eta} \frac{\partial \eta}{\partial y} \right) \left(\frac{\partial \overline{v}}{\partial \xi} \frac{\partial \xi}{\partial y} + \frac{\partial \overline{v}}{\partial \eta} \frac{\partial \eta}{\partial y} \right) \\
&= \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \xi} \left(\frac{\partial \xi}{\partial y} \frac{\partial \xi}{\partial y} \right) + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \eta} \left(\frac{\partial \eta}{\partial y} \frac{\partial \eta}{\partial y} \right) + \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \eta} \left(\frac{\partial \xi}{\partial y} \frac{\partial \eta}{\partial y} \right) + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \xi} \left(\frac{\partial \xi}{\partial y} \frac{\partial \eta}{\partial y} \right) \\
&= \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \xi} \mathcal{G}^{\xi\xi}_{yy} + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \eta} \mathcal{G}^{\eta\eta}_{yy} + \frac{\partial u}{\partial \xi} \frac{\partial \overline{v}}{\partial \eta} \mathcal{G}^{\xi\eta}_{yy} + \frac{\partial u}{\partial \eta} \frac{\partial \overline{v}}{\partial \xi} \mathcal{G}^{\xi\eta}_{yy}, \tag{3.8}
\end{aligned}
$$

introducing the short notations, $\mathcal{G}^{\xi\xi}_{xx}$, $\mathcal{G}^{\eta\eta}_{xx}$, $\mathcal{G}^{\xi\eta}_{xx}$, $\mathcal{G}^{\xi\xi}_{yy}$, $\mathcal{G}^{\eta\eta}_{yy}$, and $\mathcal{G}^{\xi\eta}_{yy}$, for the geometric factors. Using the expansion (3.1) for $u, v \in V_N$, we derive the discrete operator for (3.6) including (3.7)–(3.8) as

$$
\mathcal{A}^N(u, v) = \sum_{e=1}^E \sum_{\hat{i},\hat{j}=0}^N \sum_{i,j=0}^N v^e_{\hat{i}\hat{j}} \left(\int_I \frac{\partial \psi_{ij}}{\partial \xi} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \xi} \bar{\mathcal{G}}^{11} J d\mathbf{r} + \frac{\partial \psi_{ij}}{\partial \xi} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \eta} \bar{\mathcal{G}}^{12} J d\mathbf{r} \right) u^e_{ij} \tag{3.9}
$$

$$
+ \sum_{e=1}^E \sum_{\hat{i},\hat{j}=0}^N \sum_{i,j=0}^N v^e_{\hat{i}\hat{j}} \left(\int_I \frac{\partial \psi_{ij}}{\partial \eta} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \xi} \bar{\mathcal{G}}^{21} J d\mathbf{r} + \frac{\partial \psi_{ij}}{\partial \eta} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \eta} \bar{\mathcal{G}}^{22} J d\mathbf{r} \right) u^e_{ij}, \tag{3.10}
$$

where $Jd\mathbf{r} = Jd\xi d\eta$. On each local element, the Jacobian J and the geometric factors, defined by

$$\bar{\mathcal{G}}^{11} = (\mathcal{G}^{\xi\xi}_{xx} + \mathcal{G}^{\xi\xi}_{yy}), \qquad \bar{\mathcal{G}}^{12} = (\mathcal{G}^{\xi\eta}_{xx} + \mathcal{G}^{\xi\eta}_{yy}), \tag{3.11}$$

$$\bar{\mathcal{G}}^{21} = (\mathcal{G}^{\xi\eta}_{xx} + \mathcal{G}^{\xi\eta}_{yy}), \qquad \bar{\mathcal{G}}^{22} = (\mathcal{G}^{\eta\eta}_{xx} + \mathcal{G}^{\eta\eta}_{yy}), \tag{3.12}$$

are introduced from the coordinate transformation and computed from the following relation:

$$ J = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{vmatrix} \quad \text{from} \quad \begin{pmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{pmatrix} \begin{pmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \eta}{\partial x} & \frac{\partial \eta}{\partial y} \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.13} $$

We apply the numerical quadrature on the GLL nodes for the integrations in Eqs. (3.9)–(3.10) as

$$ \int_I \frac{\partial \psi_{ij}}{\partial \xi} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \xi} \bar{\mathcal{G}}^{11} d\mathbf{r} = \sum_{k,m=0}^N \bar{\mathcal{G}}^{11}_{km} J_{km} w_k w_m l'_i(\xi_k) l_j(\eta_m) l'_{\hat{i}}(\xi_k) l_{\hat{j}}(\eta_m), \tag{3.14} $$

$$ \int_I \frac{\partial \psi_{ij}}{\partial \xi} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \eta} \bar{\mathcal{G}}^{12} d\mathbf{r} = \sum_{k,m=0}^N \bar{\mathcal{G}}^{12}_{km} J_{km} w_k w_m l'_i(\xi_k) l_j(\eta_m) l_{\hat{i}}(\xi_k) l'_{\hat{j}}(\eta_m), \tag{3.15} $$

$$ \int_I \frac{\partial \psi_{ij}}{\partial \eta} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \xi} \bar{\mathcal{G}}^{21} d\mathbf{r} = \sum_{k,m=0}^N \bar{\mathcal{G}}^{21}_{km} J_{km} w_k w_m l_i(\xi_k) l'_j(\eta_m) l'_{\hat{i}}(\xi_k) l_{\hat{j}}(\eta_m), \tag{3.16} $$

$$ \int_I \frac{\partial \psi_{ij}}{\partial \eta} \frac{\partial \psi_{\hat{i}\hat{j}}}{\partial \eta} \bar{\mathcal{G}}^{22} d\mathbf{r} = \sum_{k,m=0}^N \bar{\mathcal{G}}^{22}_{km} J_{km} w_k w_m l_i(\xi_k) l'_j(\eta_m) l_{\hat{i}}(\xi_k) l'_{\hat{j}}(\eta_m), \tag{3.17} $$

where $\bar{\mathcal{G}}^{(\cdot)}_{km}$ and J_{km} represent the geometric values and the Jacobian at the nodal points, respectively, and w_k and w_m are the one-dimensional GLL quadrature weights. Note that $\bar{\mathcal{G}}^{12}_{km} = \bar{\mathcal{G}}^{21}_{km}$. We now have (3.6) in a discrete form as the following:

$$ \mathcal{A}^N(u,v) = \sum_{e=1}^E (v^e)^T \begin{bmatrix} \mathbf{D}_\xi \\ \mathbf{D}_\eta \end{bmatrix}^T \begin{bmatrix} \mathbf{G}^{11} & \mathbf{G}^{12} \\ \mathbf{G}^{21} & \mathbf{G}^{22} \end{bmatrix}^e \begin{bmatrix} \mathbf{D}_\xi \\ \mathbf{D}_\eta \end{bmatrix} u^e \tag{3.18} $$

$$ = \sum_{e=1}^E (v^e)^T \mathbf{D}^T \mathbf{G}^e \mathbf{D} u^e = \sum_{e=1}^E (v^e)^T \mathbf{A}^e u^e, \tag{3.19} $$

where the differentiation matrices with respect to ξ and η, $\mathbf{D}_\xi$ and $\mathbf{D}_\eta$, respectively, are written as

$$ \mathbf{D}_\xi = \mathbf{I} \otimes \hat{\mathbf{D}} \quad \text{and} \quad \mathbf{D}_\eta = \hat{\mathbf{D}} \otimes \mathbf{I} \tag{3.20} $$

in a tensor product form of the one-dimensional differentiation matrix $\hat{\mathbf{D}}_{ki} = l'_i(\xi_k)$ and the identity matrix $\mathbf{I}$ in $R^{(N+1)\times(N+1)}$. The entries of the one-dimensional differentiation matrix are

$$ \hat{\mathbf{D}}_{ij} = \frac{L_N(\xi_i)}{L_N(\xi_j)(\xi_i - \xi_j)} \;\; (i \neq j); \;\; \hat{\mathbf{D}}_{00} = -\frac{(N+1)N}{4}; \;\; \hat{\mathbf{D}}_{NN} = \frac{(N+1)N}{4}; \;\; \hat{\mathbf{D}}_{ii} = 0, \tag{3.21} $$

which is skew-centrosymmetric $\hat{\mathbf{D}}_{ij} = -\hat{\mathbf{D}}_{N-i,N-j}$. Equation (3.18) involves the pointwise multiplication of the nodal values $u^e = [u^e_l]$ by each diagonal component of $\mathbf{G}^{(\cdot)} = [\mathbf{G}^{(\cdot)}_l] = \text{diag}\{\bar{\mathcal{G}}^{(\cdot)}_{km} J_{km} w_k w_m\}$ for $l = k + (N+1)(m-1)$ on the nodal points on each local element Ω^e. Let us denote the stiffness matrix on Ω as $\mathbf{A}$, using the local stiffness matrices $\mathbf{A}^e$, represented by

$$ \mathbf{A} = \begin{bmatrix} \mathbf{A}^1 & & & & \\ & \ddots & & & \\ & & \mathbf{A}^e & & \\ & & & \ddots & \\ & & & & \mathbf{A}^E \end{bmatrix} \quad \text{with} \quad \mathbf{A}^e = \mathbf{D}^T \mathbf{G}^e \mathbf{D}. \tag{3.22} $$

Then we can write Eq. (3.18) simply as

$$\mathcal{A}^N(u, v) = \mathbf{v}^T \mathbf{A} \mathbf{u}. \tag{3.23}$$

Here we note that the matrix $\mathbf{A}$ is symmetric from the fact that

$$(\mathbf{A}^e)^T = (\mathbf{D}^T \mathbf{G}^e \mathbf{D})^T = \mathbf{D}^T (\mathbf{G}^e)^T \mathbf{D} = \mathbf{D}^T (\mathbf{G}^e) \mathbf{D} = \mathbf{A}^e. \tag{3.24}$$

Arithmetic Operations: The matrix $\mathbf{A}$ is never explicitly formed. We perform matrix-matrix multiplication acting only on the block diagonal matrices $\mathbf{A}^e$. We begin with the tensor product based derivative evaluations (3.18) that can be recast as matrix-matrix products on each element:

$$\mathbf{u}_\xi := (\mathbf{I} \otimes \hat{\mathbf{D}}) u^e := \hat{\mathbf{D}}[u]^e, \tag{3.25}$$

$$\mathbf{u}_\eta := (\hat{\mathbf{D}} \otimes \mathbf{I}) u^e := [u]^e \hat{\mathbf{D}}^T, \tag{3.26}$$

where u^e is a vector arranged in columnwise consecutive entries of u_{ij}^e, advancing with the leading index (i) as shown in (3.4). In (3.25), u^e is treated as an $(N+1) \times (N+1)$ matrix, denoted by $[u]^e$ as

$$[u]^e = \begin{bmatrix} u_{00}^e & u_{01}^e & \cdots & u_{0N}^e \\ \vdots & \vdots & \ddots & \vdots \\ u_{N0}^e & u_{N1}^e & \cdots & u_{NN}^e \end{bmatrix}. \tag{3.27}$$

This requires $2E(N+1)^3$ operations on Ω. The pointwise multiplications with the geometric factors $\mathbf{u}_x = \mathbf{G}^{11}\mathbf{u}_\xi + \mathbf{G}^{12}\mathbf{u}_\eta$ and $\mathbf{u}_y = \mathbf{G}^{21}\mathbf{u}_\xi + \mathbf{G}^{22}\mathbf{u}_\eta$ require $6E(N+1)^2$ operations. Then we compute the summation of transposed derivative operators, $\mathbf{D}_\xi \mathbf{u}_x + \mathbf{D}_\eta \mathbf{u}_y$, involving $4E(N+1)^3 + E(N+1)^2$ operations. Thus the total operation for $\mathbf{A}\mathbf{u}$ is $6E(N+1)^3 + 7E(N+1)^2$. The leading-order storage requirement for the factored stiffness matrix is $3E(N+1)^2$, because of the relation $\mathbf{G}_{12} = \mathbf{G}_{21}$ on Ω^e.

Direct Stiffness Summation: The solution vector in (3.23) is based on the unassembled representation, recalling Figure 2(a), without applying the continuity at the element interface between neighboring elements. To construct the solution vector to be continuous across element interfaces on the coincident nodal values,

$$(x_{ij}, y_{ij})^e = (x_{\hat{i}\hat{j}}, y_{\hat{i}\hat{j}})^{\hat{e}} \rightarrow u_{ij}^e = u_{\hat{i}\hat{j}}^{\hat{e}} \quad \text{for} \quad e \neq \hat{e}, \tag{3.28}$$

we introduce a Boolean connectivity matrix $\mathbf{Q}$ [7] that maps the global representation $\underline{\mathbf{u}}$ to the local representation $\mathbf{u}$, and its transpose $\mathbf{Q}^T$ that maps the local representation $\mathbf{u}$ to the global representation $\underline{\mathbf{u}}$. Then we can define the following:

$$\mathbf{u} = \mathbf{Q}\underline{\mathbf{u}} \quad \text{and} \quad \mathbf{u}^* = \mathbf{Q}^T \mathbf{u}. \tag{3.29}$$

The action of $\mathbf{Q}$ on $\underline{\mathbf{u}}$ returns the copy entries of $\underline{\mathbf{u}}$ on the coincident nodes, referred as the scatter operation. The action of $\mathbf{Q}^T$ on $\mathbf{u}$ returns $\mathbf{u}^*$ with the sum entries of $\mathbf{u}$ on the coincident nodes, referred as the gather operation. The interior nodes are unchanged from both of the actions. Using these matrices, we can rewrite Eq. (3.23) for the continuous solution $\underline{\mathbf{u}}$ as

$$\mathcal{A}^N(u, v) = \underline{\mathbf{v}}^T \mathbf{Q}^T \mathbf{A} \mathbf{Q} \underline{\mathbf{u}} = \underline{\mathbf{v}}^T \bar{\mathbf{A}} \underline{\mathbf{u}}. \tag{3.30}$$

For a continuous solution $\mathbf{u}$ in the local ordering representation, the following equivalence holds:

$$\mathbf{Q}^T \mathbf{A} \mathbf{Q} \underline{\mathbf{u}} \iff (\mathbf{Q}\mathbf{Q}^T) \mathbf{A} \mathbf{u}. \tag{3.31}$$

11

We note that the gather-scatter operation $\mathbf{Q}\mathbf{Q}^T$ can be viewed as a single operation, involving summation of the variables on the shared interface nodes and redistribution of them to their original locations within one communication. The operation is referred as *direct stiffness summation*, or simply *dssum*. In this paper, we use the following notation for the gather-scatter operation:

$$dssum := \mathbf{Q}\mathbf{Q}^T. \tag{3.32}$$

In practical implementations, we write our algorithms in an element-based format by utilizing matrix-vector products evaluated independently on each local element. Thus it is natural to consider the *dssum* approach and perform the local-to-local transformation as in the right-hand side of (3.31), that is, $dssum(\mathbf{Au})$. We build a local-to-global mapping array to handle the actions of $\mathbf{Q}$ and $\mathbf{Q}^T$ without constructing $\mathbf{Q}$ and $\mathbf{Q}^T$ explicitly. A detailed description of the algorithms and parallel implementations can be found in Chapter 4 and Chapter 8 of [7].

3.2 Mass Matrices

To obtain the mass matrix, we consider the following inner product:

$$\mathcal{B}(u,v) = \int_{\Omega} u\overline{v}d\Omega, \tag{3.33}$$

which can be discretized as

$$
\begin{aligned}
\mathcal{B}^N(u,v) &= \sum_{e=1}^{E}\sum_{\hat{i},\hat{j}=0}^{N}\sum_{i,j=0}^{N} v_{ij}^e \left(\int_{\Omega^e} \psi_{ij}\psi_{\hat{i}\hat{j}}d\Omega \right) u_{ij}^e = \sum_{e=1}^{E}\sum_{\hat{i},\hat{j}=0}^{N}\sum_{i,j=0}^{N} v_{ij}^e \left(\int_{I} \psi_{ij}\psi_{\hat{i}\hat{j}}Jd\mathbf{r} \right) u_{ij}^e \\
&= \sum_{e=1}^{E}\sum_{\hat{i},\hat{j}=0}^{N}\sum_{i,j=0}^{N} v_{ij}^e \left(\sum_{k,m=0}^{N} J_{km}w_k w_m l_i(\xi_k)l_j(\eta_m)l_{\hat{i}}(\xi_k)l_{\hat{j}}(\eta_m) \right) u_{ij}^e \\
&= \sum_{e=1}^{E}(v^e)^T \mathbf{J}^e \left(\hat{\mathbf{M}} \otimes \hat{\mathbf{M}} \right) u^e = \sum_{e=1}^{E}(v^e)^T \mathbf{B}^e u^e, \tag{3.34}
\end{aligned}
$$

where $\hat{\mathbf{M}} = \mathrm{diag}\{w_k\}$ is the one-dimensional mass matrix and $\mathbf{J}^e = [\mathbf{J}_{ll}^e] = \mathrm{diag}\{J_{km}\}$ for $l = k + (N+1)(m-1)$. We can denote the mass matrix $\mathbf{B}$, using the local mass matrix $\mathbf{B}^e$ as

$$
\mathbf{B} = \begin{bmatrix} \mathbf{B}^1 & & & & \\ & \ddots & & & \\ & & \mathbf{B}^e & & \\ & & & \ddots & \\ & & & & \mathbf{B}^E \end{bmatrix} \quad \text{with} \quad \mathbf{B}^e = \mathbf{J}^e(\hat{\mathbf{M}} \otimes \hat{\mathbf{M}}), \tag{3.35}
$$

which is fully diagonal. Then we can write Eq. (3.34) simply as

$$\mathcal{B}^N(u,v) = \mathbf{v}^T \mathbf{B}\mathbf{u}. \tag{3.36}$$

For continuous solution, Eq. (3.36) in the assembled representation can be expressed as

$$\mathcal{B}^N(u,v) = \underline{\mathbf{v}}^T \mathbf{Q}^T \mathbf{B}\mathbf{Q}\underline{\mathbf{u}} = \underline{\mathbf{v}}^T \bar{\mathbf{B}}\underline{\mathbf{u}}. \tag{3.37}$$

12

3.3 Quasi-Periodic Matrix

We consider the following inner product for the quasi-periodic operator in Eq. (2.35):

$$C(u,v) = \int_\Omega \frac{\partial u}{\partial x}\overline{v}d\Omega, \tag{3.38}$$

which can be discretized as

$$
\begin{aligned}
\mathcal{C}^N(u,v) &= \sum_{e=1}^{E}\sum_{\hat{i},\hat{j}=0}^{N}\sum_{i,j=0}^{N} v_{ij}^e \left(\int_{\Omega^e}\frac{\partial\psi_{ij}}{\partial x}\psi_{\hat{i}\hat{j}}d\Omega\right)u_{ij}^e = \sum_{e=1}^{E}\sum_{\hat{i},\hat{j}=0}^{N}\sum_{i,j=0}^{N} v_{ij}^e \left(\int_{I}\frac{\partial\psi_{ij}}{\partial x}\psi_{\hat{i}\hat{j}}Jd\mathbf{r}\right)u_{ij}^e \\
&= \sum_{e=1}^{E}\sum_{\hat{i},\hat{j}=0}^{N}\sum_{i,j=0}^{N} v_{ij}^e \left(\sum_{k,m=0}^{N} J_{km}w_k w_m l_i'(\xi_k)l_j(\eta_m)l_{\hat{i}}(\xi_k)l_{\hat{j}}(\eta_m)\right)u_{ij}^e \\
&= \sum_{e=1}^{E}(v^e)^T\mathbf{J}^e\left(\hat{\mathbf{M}}\otimes\hat{\mathbf{M}}\hat{\mathbf{D}}\right)u^e = \sum_{e=1}^{E}(v^e)^T\mathbf{J}^e\left(\hat{\mathbf{M}}\otimes\hat{\mathbf{C}}\right)u^e = \sum_{e=1}^{E}(v^e)^T\mathbf{C}^e u^e. \tag{3.39}
\end{aligned}
$$

By convention, (3.38) could be referred as a convective operator in computational fluids. In this context, this relates to the quasi-periodic term in (2.14) for the oblique incident ($\alpha\neq 0$). Thus we refer to it as a quasi-periodic operator, because the operator is a derivative, resulting from imposing the periodicity for the quasi-periodic solution in (2.12). Then, we define quasi-periodic matrix $\mathbf{C}$ on Ω using the local quasi-periodic matrices as

$$
\mathbf{C} = \begin{bmatrix} \mathbf{C}^1 & & & & \\ & \ddots & & & \\ & & \mathbf{C}^e & & \\ & & & \ddots & \\ & & & & \mathbf{C}^E \end{bmatrix} \quad \text{with} \quad \mathbf{C}^e = \mathbf{J}^e(\hat{\mathbf{M}}\otimes\hat{\mathbf{C}}). \tag{3.40}
$$

We can write Eq. (3.39) simply as

$$\mathcal{C}^N(u,v) = \mathbf{v}^T\mathbf{C}\mathbf{u}. \tag{3.41}$$

For the continuous solution, Eq. (3.41) in the assembled representation can be expressed as

$$\mathcal{C}^N(u,v) = \underline{\mathbf{v}}^T\mathbf{Q}^T\mathbf{C}\mathbf{Q}\underline{\mathbf{u}} = \underline{\mathbf{v}}^T\bar{\mathbf{C}}\underline{\mathbf{u}}. \tag{3.42}$$

3.4 Spectral Element Dirichlet-to-Neumann Boundary Operator

In this section, we formulate a spectral element discretization for the transparent boundary operator or simply DtN (Dirichlet-to-Neumann) boundary operator T. Let us denote $\Gamma = \cup_{\hat{e}=1}^{\hat{E}}\Gamma^{\hat{e}}$, where $\Gamma^{\hat{e}} = \Omega^{\hat{e}}\cap\partial\Omega$ are nonoverlapping DtN boundary surfaces on the local elements $\Omega^{\hat{e}}$. We define a dtn-to-local mapping array that contains the indices of the DtN surface nodes $(i,j,\hat{e})$ to the local index $(i,j,e) := \text{dtn-to-local}(i,j,\hat{e})$. We note that the DtN boundary nodes in y fall on the index either with $j=0$ or with $j=N$, which will be represented simply by a fixed index as $j=j_b$.

DtN Matrix T: We can represent our approximate solution on $\Gamma^{\hat{e}}$ in the form of (3.1) as

$$u^{\hat{e}}(x,b) = \sum_{i,j=0}^{N} u_{ij}^{\hat{e}}l_i(\xi)l_j(\eta(b)) = \sum_{i=0}^{N} u_{ij_b}^{\hat{e}}l_i(\xi). \tag{3.43}$$

13

From Eqs. (2.43)–(2.44), we have

$$T(u,v) = \int_\Gamma T[u]\overline{v}d\Gamma = \int_\Gamma \left(\sum_{p=-\infty}^{\infty} \mathbf{i}\beta_p \hat{u}_p e^{\mathbf{i}d_p x} \right) \overline{v}dx = \sum_{p=-\infty}^{\infty} \mathbf{i}\beta_p \hat{u}_p \int_\Gamma e^{\mathbf{i}d_p x}\overline{v}dx, \qquad (3.44)$$

where $d_p = \frac{2\pi p}{d}$ and $\hat{u}_p$ are the one-dimensional Fourier coefficients of $u(x,b)$ on Γ given as

$$\hat{u}_p = \frac{1}{L}\int_0^L u(x',b)e^{-\mathbf{i}d_p x'}dx' \approx \frac{1}{L}\sum_{\hat{e}=1}^{\hat{E}} \int_{\Gamma^{\hat{e}}} u^{\hat{e}}(x',b)e^{-\mathbf{i}d_p x'}dx'. \qquad (3.45)$$

Plugging (3.45) into (3.44) with a finite expansion of $T[u]$ ($|p| \le P$) and applying (3.43), we have

$$\begin{aligned}
T^N(u,v) &= \sum_{p=-P}^{P} \mathbf{i}\beta_p \left(\frac{1}{L}\sum_{\hat{e}=1}^{\hat{E}} \int_{\Gamma^{\hat{e}}} u^{\hat{e}}(x',b)e^{-\mathbf{i}d_p x'}dx' \right) \left(\sum_{\overline{e}=1}^{\hat{E}} \int_{\Gamma^{\overline{e}}} e^{\mathbf{i}d_p x}\overline{v}dx \right) && (3.46) \\
&= \sum_{p=-P}^{P} \mathbf{i}\beta_p \left[\sum_{i=0}^{N} \left(\frac{1}{L}\sum_{\hat{e}=1}^{\hat{E}} u^{\hat{e}}_{ij_b} \int_{\Gamma^{\hat{e}}} l_i(\xi)e^{-\mathbf{i}d_p x'}dx' \right) \right] \left(\sum_{\overline{e}=1}^{\hat{E}} \int_{\Gamma^{\overline{e}}} e^{\mathbf{i}d_p x}\overline{v}dx \right). && (3.47)
\end{aligned}$$

Choosing $\overline{v} = l_{\hat{i}}(\xi)$ with a different index set of $\hat{i}$ on each $\Omega^{\hat{e}}$ and defining the following,

$$s_i^{\hat{e},p} = \frac{1}{\sqrt{L}}\int_{\Gamma^{\hat{e}}} l_i(\xi)e^{-\mathbf{i}d_p x'}\, dx' \quad \text{and} \quad s_{\hat{i}}^{\overline{e},-p} = \frac{1}{\sqrt{L}}\int_{\Gamma^{\overline{e}}} l_{\hat{i}}(\xi)e^{\mathbf{i}d_p x}\, dx, \qquad (3.48)$$

we can express (3.44) in a simplified form as

$$T^N(u,v) = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b} \left[\sum_{p=-P}^{P} \mathbf{i}\beta_p \left(s_i^{\hat{e},p} \right) \left(\sum_{\overline{e}=1}^{\hat{E}} s_{\hat{i}}^{\overline{e},-p} \right) \right] = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b} T^{\hat{e}}_{\hat{i}i}. \qquad (3.49)$$

Here we note that $s_i^{\hat{e},-p}$ is the complex conjugate of $s_i^{\hat{e},p}$ from the following:

$$\overline{s_i^{\hat{e},p}} = \overline{\frac{1}{\sqrt{L}}\int_{\Gamma^{\hat{e}}} l_i(\xi)e^{-\mathbf{i}d_p x}\, dx} = \frac{1}{\sqrt{L}}\int_{\Gamma^{\hat{e}}} l_i(\xi)e^{\mathbf{i}d_p x}\, dx = s_i^{\hat{e},-p}. \qquad (3.50)$$

Thus we need only to compute $s_i^{\hat{e},p}$ for $p \ge 0$ to obtain

$$T^{\hat{e}}_{\hat{i}i} = \mathbf{i}\left(\beta_0 s_i^{\hat{e},0} \sum_{\overline{e}=1}^{\hat{E}} s_{\hat{i}}^{\overline{e},0} + \sum_{p=1}^{P} \left[\beta_p s_i^{\hat{e},p} + \beta_{-p}\overline{s_i^{\hat{e},p}} \right] \sum_{\overline{e}=1}^{\hat{E}} \overline{s_{\hat{i}}^{\overline{e},p}} \right), \qquad (3.51)$$

where $\beta_p = \beta_{-p}$ only if $\alpha = 0\mathbf{i}$; $\beta_p \ne \beta_{-p}$ for $\alpha \ne 0$. Therefore, no particular relation can be found between β_p and β_{-p} in general. Here $T^{\hat{e}}_{\hat{i}i}$ is a complex number, so we can alternatively write (3.49) as

$$T^N(u,v) = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b} T^{\hat{e}}_{\hat{i}i} = \sum_{\hat{e}=1}^{\hat{E}}\sum_{i=0}^{N} u^{\hat{e}}_{ij_b} \left[(T^{\hat{e}}_{\hat{i}i})_{\text{real}} + \mathbf{i}(T^{\hat{e}}_{\hat{i}i})_{\text{imag}} \right]. \qquad (3.52)$$

Now, we can map the values of $T^{\hat{e}}_{\hat{i}i}$ into a matrix $\mathbf{T}^e = \left[\mathbf{T}^e_{\hat{l}l} \right]$ for $\hat{l} = \hat{i} + (N+1)j$ and $l = i + (N+1)j$ from the dtn-to-local mapping $(\hat{i}, j, e) := \text{dtn-to-local}(\hat{i}, j_b, \hat{e})$ and $(i, j, e) := \text{dtn-to-local}(i, j_b, \hat{e})$.

Similarly, $\{u^{\hat{e}}_{ij_b}\}$ can be mapped to the local data $\{u^e_{ij}\}$. Note that the entries of $\mathbf{T}^e$ are zeros if the indices are not indicating the DtN boundary nodes. We now have Eq. (3.49) in the local representation form as

$$\mathcal{T}^N(u,v) = \sum_{e=1}^{E}(v^e)^T\mathbf{T}^e u^e = \mathbf{v}^T\mathbf{T}\mathbf{u} = \mathbf{v}^T(\mathbf{T_r} + \mathbf{i}\mathbf{T_i})\mathbf{u}, \tag{3.53}$$

where $\mathbf{T_r}$ and $\mathbf{T_i}$ represent the real and imaginary part of the complex matrix $\mathbf{T}$. Thus we have the assembled representation of (3.53) as

$$\mathcal{T}^N(u,v) = \underline{\mathbf{v}}^T\mathbf{Q}^T\mathbf{T}\mathbf{Q}\underline{\mathbf{u}} = \underline{\mathbf{v}}^T\bar{\mathbf{T}}\underline{\mathbf{u}} = \underline{\mathbf{v}}^T(\bar{\mathbf{T}}_r + \mathbf{i}\bar{\mathbf{T}}_i)\underline{\mathbf{u}}. \tag{3.54}$$

For ρ in (2.43), we apply notations similar to those used for u. Then we have the following:

$$\mathcal{F}^N(\rho,v) = \sum_{e=1}^{E}(v^e)^T\mathbf{B}^e\rho^e = \mathbf{v}^T\mathbf{B}\boldsymbol{\rho} = \mathbf{v}^T\mathbf{F}\boldsymbol{\rho}, \tag{3.55}$$

with the assembled representation as

$$\mathcal{F}^N(\rho,v) = \underline{\mathbf{v}}^T\mathbf{Q}^T\mathbf{B}\mathbf{Q}\underline{\boldsymbol{\rho}} = \underline{\mathbf{v}}^T\bar{\mathbf{F}}\underline{\boldsymbol{\rho}}. \tag{3.56}$$

Compute Matrix T: We next discuss how to compute $s_i^{\hat{e},p}$ in Eq. (3.51). Note that the data is precomputed only one time. One might apply the GLL quadrature for the integrations when d_p is small. For large d_p, however, the GLL quadrature is not accurate enough to capture the high-frequency modes. One can consider the discrete FFT algorithm since it is the pth component of the inverse DFFT of function $l_i(\xi)$. However, $l_i(\xi)$ has only a very small portion of compact support on Γ, so we can compute it directly on its local compact support using refined GLL quadrature points on each $\Gamma^{\hat{e}}$. Another approach is to use the relation to the Bessel function, which can be more efficient than the other approach.

In this paper, we discuss the computation of $s_i^{\hat{e},p}$ based on the Bessel function representation. We have written $l_i(\xi)$ in the finite expansion of the mth-order Legendre polynomials given as

$$l_i(\xi) = \sum_{m=0}^{N}(\hat{l}_i)_m L_m(\xi), \tag{3.57}$$

where $(\hat{l}_i)_m$ are the Legendre expansion coefficients defined by

$$(\hat{l}_i)_m = \frac{2m+1}{2}\int_{-1}^{1}l_i(\xi)L_m(\xi)d\xi. \tag{3.58}$$

Then, substituting (3.57) in (3.48) and using simply the notation x, instead of x', we have

$$s_i^{\hat{e},p} = \frac{1}{\sqrt{L}}\int_{\Gamma^{\hat{e}}}l_i(\xi(x))e^{-\mathbf{i}d_p x}dx = \frac{1}{\sqrt{L}}\sum_{m=0}^{N}(\hat{l}_i)_m\left(\int_{-1}^{1}L_m(\xi)e^{-\mathbf{i}d_p x(\xi)}J_s^{\hat{e}}d\xi\right), \tag{3.59}$$

where $J_s^{\hat{e}}$ is the surface Jacobian on $\Gamma^{\hat{e}}$. In fact, each $\Gamma^{\hat{e}}$ is represented by an interval $[x_{\min}^{\hat{e}}, x_{\max}^{\hat{e}}]$ with the coordinate transformation by $x(\xi) = \hat{a}_e\xi + \hat{b}_e$ with $\hat{a}_e = (x_{\max}^{\hat{e}} - x_{\min}^{\hat{e}})/2$ and $\hat{b}_e = (x_{\max}^{\hat{e}} + x_{\min}^{\hat{e}})/2$, so that $J_s^{\hat{e}} \equiv \hat{a}_e$ is constant on $\Gamma^{\hat{e}}$. Then, Eq. (3.59) becomes

$$s_i^{\hat{e},p} = \frac{\hat{a}_e}{\sqrt{L}}\sum_{m=0}^{N}(\hat{l}_i)_m d_m^{p,\hat{e}} \quad \text{with} \quad q_m^{p,\hat{e}} = \int_{-1}^{1}L_m(\xi)e^{-\mathbf{i}d_p(\hat{a}_e\xi + \hat{b}_e)}d\xi. \tag{3.60}$$

15

Now we need to compute the two terms $(\hat{l}_i)_m$ and $q_m^{p,\hat{e}}$, in (3.60). To compute $(\hat{l}_i)_m$, one might apply the GLL quadrature for the integration term in (3.58) as follows:

$$(\hat{l}_i)_m = \frac{2m+1}{2} \sum_{k=0}^{N} l_i(\xi_k) L_m(\xi_k) w_k = \frac{2m+1}{2} L_m(\xi_i) w_i. \tag{3.61}$$

An alternative approach is to evaluate (3.57) on the GLL grids in $[-1, 1]$, resulting in the form

$$\mathbf{L}\hat{\mathbf{L}} = \begin{bmatrix} L_0(\xi_0) & L_1(\xi_0) & \cdots & L_m(\xi_0) \\ \vdots & \vdots & \vdots & \vdots \\ L_0(\xi_N) & L_1(\xi_N) & \cdots & L_m(\xi_N) \end{bmatrix} \begin{bmatrix} (\hat{l}_0)_0 & \cdots (\hat{l}_i)_0 & \cdots & (\hat{l}_N)_0 \\ \vdots & \vdots & \vdots & \vdots \\ (\hat{l}_0)_N & \cdots (\hat{l}_i)_N & \cdots & (\hat{l}_N)_N \end{bmatrix} \equiv \mathbf{I}, \tag{3.62}$$

and compute the inverse of the matrix $\mathbf{L} = [\mathbf{L}_{ji}] = [L_i(\xi_j)]$ to obtain $\hat{\mathbf{L}} = [\hat{\mathbf{L}}_{mi}] = [(\hat{l}_i)_m] = \mathbf{L}^{-1}$. To compute $q_m^{p,\hat{e}}$, we recall that the Legendre polynomials are related to the Bessel functions as

$$\int_{-1}^{1} L_m(\xi) e^{-\mathbf{i}x\xi} d\xi = \frac{1}{\mathbf{i}^m} \sqrt{\frac{2\pi}{x}} J_{m+1/2}(x) = \frac{2}{\mathbf{i}^m} j_m(x) \quad \text{for} \quad x \in R, \tag{3.63}$$

where j_m is the spherical Bessel function and J_m is the ordinary Bessel function with the relation

$$j_m(x) = \sqrt{\frac{\pi}{2x}} J_{m+1/2}(x). \tag{3.64}$$

Then, we can write

$$q_m^{p,\hat{e}} = \int_{-1}^{1} L_m(\xi) e^{-\mathbf{i}d_p(\hat{a}_e\xi + \hat{b}_e)} d\xi = e^{-\mathbf{i}d_p\hat{b}_e} \left(\frac{2}{\mathbf{i}^m} j_m(d_p\hat{a}_e) \right). \tag{3.65}$$

From (3.61) and (3.65), we have the final form of $s_i^{\hat{e},p}$ by

$$s_i^{\hat{e},p} = \frac{\hat{a}_e e^{-\mathbf{i}d_p\hat{b}_e}}{\sqrt{L}} \sum_{m=0}^{N} (\hat{l}_i)_m \left(\frac{2}{\mathbf{i}^m} j_m(d_p\hat{a}_e) \right). \tag{3.66}$$

3.5 Matrix Structures and Eigenvalues

In this section, we discuss a complete set of our SE scheme including boundary conditions, provided with the matrix structures and eigenvalue distributions for the operators. We arrange our solution as a single real vector with the length of $2n$ expressed by $\mathbf{u}^N = [u_{\mathrm{r}}^N, u_{\mathrm{i}}^N]^T$ where u_{r}^N and u_{i}^N represent real and imaginary part of the solution. The SE discretization leads to a linear system:
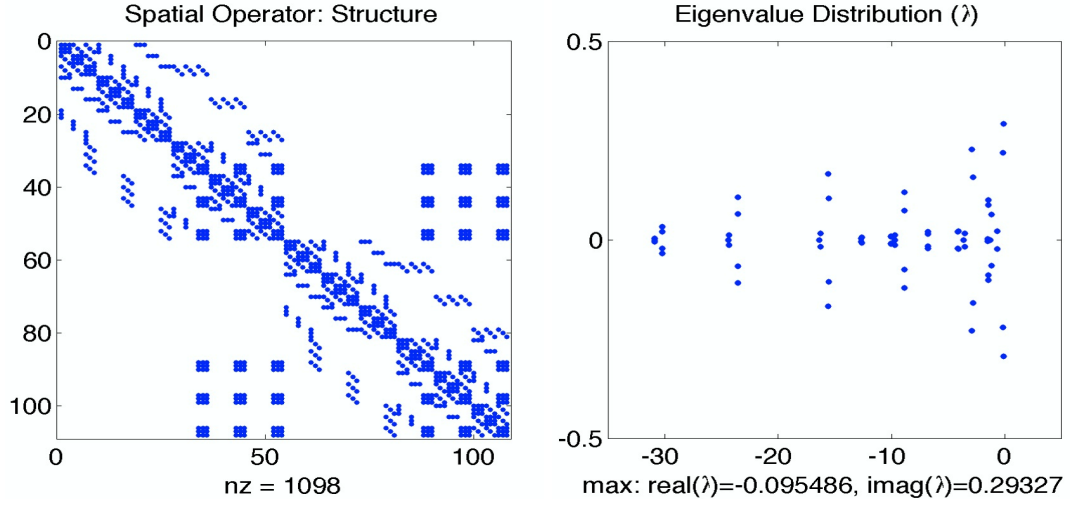
$$\mathcal{H}\mathbf{u}^N = \mathcal{F}, \tag{3.67}$$

where

$$\mathcal{H} := \begin{bmatrix} \mathbf{A} - (k^2 - \alpha^2)\mathbf{B} + \mathbf{T}_{\mathrm{r}} & -\mathbf{T}_{\mathrm{i}} - 2\alpha\mathbf{C} \\ \mathbf{T}_{\mathrm{i}} + 2\alpha\mathbf{C} & \mathbf{A} - (k^2 - \alpha^2)\mathbf{B} + \mathbf{T}_{\mathrm{r}} \end{bmatrix} \quad \text{and} \quad \mathcal{F} := \begin{bmatrix} \mathbf{F}\boldsymbol{\rho}_{\mathrm{r}} \\ \mathbf{F}\boldsymbol{\rho}_{\mathrm{i}} \end{bmatrix}. \tag{3.68}$$

The Eq. (3.67) in the assembled representation can be expressed as

$$\bar{\mathcal{H}}\underline{\mathbf{u}}^N = \bar{\mathcal{F}}, \tag{3.69}$$

Table 1: Condition numbers for $\bar{\mathcal{H}}$

DtN (top)			DtN (top/bottom)		
E	N	Condition #	E	N	Condition #
3×2	3	1.187394774856384e+02	3×2	3	5.954213579989568e+01
	5	4.600292693443866e+02		5	2.347021228844633e+02
	7	1.170405666580731e+03		7	5.902204731597798e+02
	9	2.400022447753930e+03		9	1.203134665072308e+03
	11	4.290943675185254e+03		11	2.143766952778180e+03
	13	6.985462356603306e+03		13	3.482577618583732e+03

(a) $\bar{\mathcal{H}}$ with DtN/Dirichlet (top/bottom)
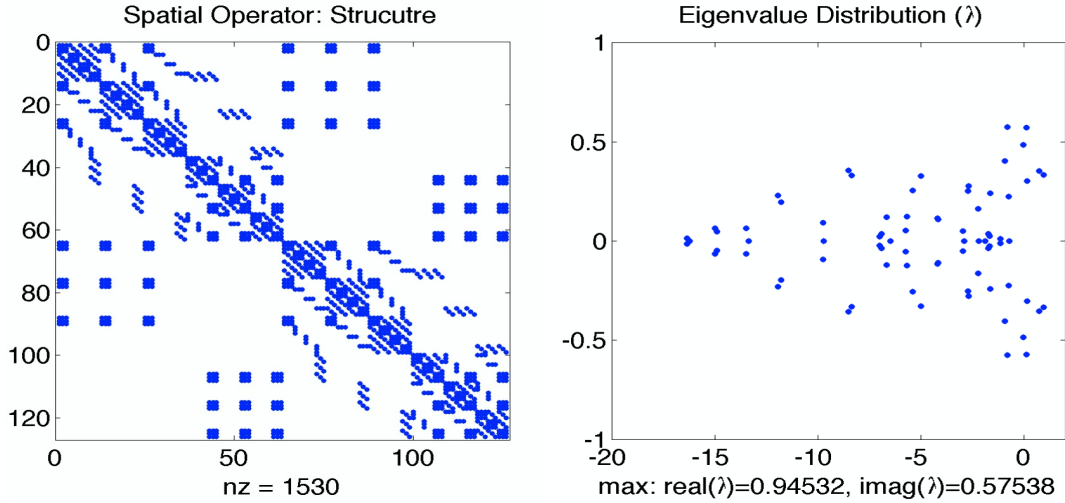
(b) $\bar{\mathcal{H}}$ with DtN/DtN (top/bottom)

Figure 3: Matrix structures (assembled): $E = 3 \times 2$ and $N = 3$.

where

$$\bar{\mathcal{H}} := \left[\begin{array}{cc} \bar{\mathbf{A}} - (k^2 - \alpha^2)\bar{\mathbf{B}} + \bar{\mathbf{T}}_{\mathrm{r}} & -\bar{\mathbf{T}}_{\mathrm{i}} - 2\alpha\bar{\mathbf{C}} \\ \bar{\mathbf{T}}_{\mathrm{i}} + 2\alpha\bar{\mathbf{C}} & \bar{\mathbf{A}} - (k^2 - \alpha^2)\bar{\mathbf{B}} + \bar{\mathbf{T}}_{\mathrm{r}} \end{array} \right] \quad \text{and} \quad \bar{\mathcal{F}} := \left[\begin{array}{c} \bar{\mathbf{F}}\underline{\rho}_{\mathrm{r}} \\ \bar{\mathbf{F}}\underline{\rho}_{\mathrm{i}} \end{array} \right]. \tag{3.70}$$

We demonstrate the matrix structures for our SE operators for different set of the boundary conditions in Figures 3(a)–3(b), provided with their eigenvalue distributions on the right panels. In Table 1, we demonstrate the condition numbers for those operators. We solve the linear system (3.69) using the generalized minimum residual (GMRES) method [24].

4 Computational Results

In this section, we examine exterior scattering problems with three types of scattering surfaces: smooth flat, smooth curved, and nonsmooth surface periodic gratings. We consider different angles of incident impinging on the scattering surface in single- and double-layer media. We solve the scalar Helmholtz equation featured with quasi-periodicity and compute the total field in a finite computational domain with transparent boundary condition based on the spectral element discretization. For validation of our computational approach, we examine the scattered field solution. For the cases of smooth flat surface periodic gratings, we compare our results with analytic solutions and provide convergence studies. For smooth curved surface periodic gratings, we consider sinusoidal grooves and compare our results with those from the TFE method [22]. For nonsmooth surface periodic gratings, we consider rectangular, triangular, and sawtooth grooves in double-layer media and demonstrate our computational solutions provided with the convergence results measured by the energy defect [4, 5, 6, 23].

4.1 Smooth Flat Scattering Surfaces

We consider single- and double-layer media with smooth flat surface periodic gratings in the x-direction. Dirichlet and DtN boundary conditions are applied in the y-direction. For these configurations, there exist analytic solutions for incident waves at arbitrary angles $\kappa = (\alpha, \beta)$.

Single Layer Consider a finite computational domain $\Omega = [0, 2\pi] \times [0, 1]$ with the scattering surface defined by $\Gamma_g = \{(x, y) \in \Omega | y = 0\}$ and transparent boundary defined on $\Gamma = \{(x, y) \in \Omega | y = 1\}$. We apply the homogeneous Dirichlet boundary condition on the soft scatterer Γ_g and the DtN boundary operator on Γ. Figure 4(a) shows our quadrilateral element mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$. Considering a given incident field $U_{\mathrm{inc}}^{\mathrm{exact}}$ impinging on Γ_g, we have the total field solution U^{exact} as

$$\begin{aligned} U_{\mathrm{inc}}^{\mathrm{exact}}(x, y) &= e^{\mathbf{i}(\alpha x - \beta(y+1))}, \tag{4.1} \\ U^{\mathrm{exact}}(x, y) &= e^{\mathbf{i}(\alpha x - \beta(y+1))} - e^{\mathbf{i}(\alpha x + \beta(y-1))}, \tag{4.2} \end{aligned}$$

where $\beta = \sqrt{k^2 - \alpha^2}$. For a fixed wavenumber $k = 1.5$ in the single layer medium, we consider incident waves for $\alpha = 0.1$ and $\alpha = 1.0$. Figures 4(b)–4(c) show the numerical solutions of the scattered fields that are obtained by subtracting the incident field from the total field: $\mathbf{U}_{\mathrm{scat}}^N = \mathbf{U}^N - \mathbf{U}_{\mathrm{inc}}^N$, where $\mathbf{U}_{\mathrm{inc}}^N$ denotes the incident field $U_{\mathrm{inc}}^{\mathrm{exact}}$ evaluated on the GLL grids.
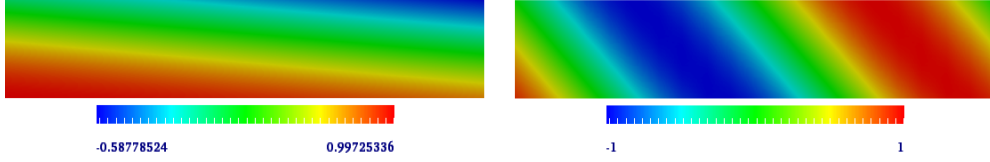
Double Layer Consider a computational domain $\Omega = [0, 2\pi] \times [-1, 1]$ with the scattering surface $\Gamma_g = \{(x, y) \in \Omega | y = 0\}$ and transparent boundaries defined on $\Gamma = \Gamma^+ \cup \Gamma^-$ for $\Gamma^+ = \{(x, y) \in \Omega | y = 1\}$ and $\Gamma^- = \{(x, y) \in \Omega | y = -1\}$. We apply the DtN boundary operator on the GLL points
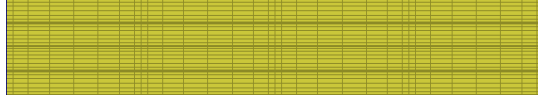
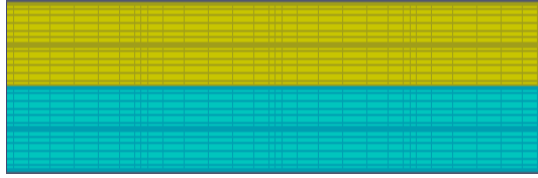(a) Mesh and GLL grids with $E = 4 \times 4$, $N = 8$
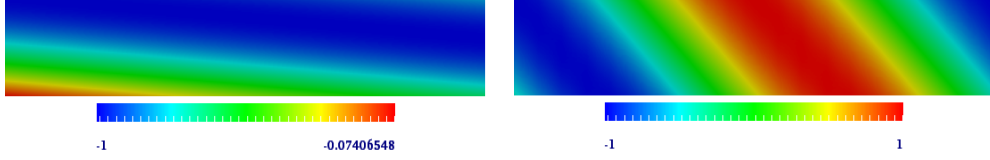
(b) Real part of the scattered field $\mathbf{U}_{\text{scat}}^N$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)
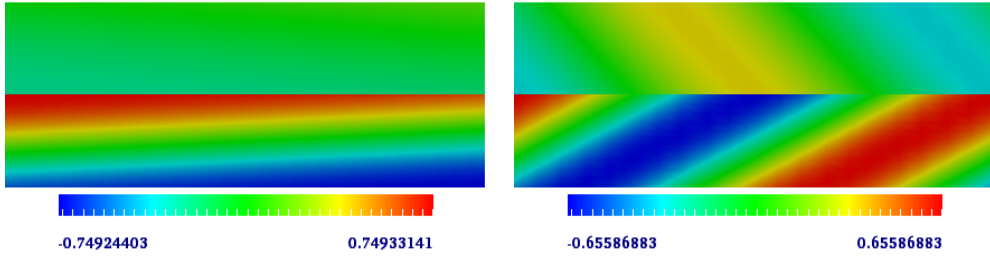

(c) Imaginary part of the scattered field $\mathbf{U}_{\text{scat}}^N$: with $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)
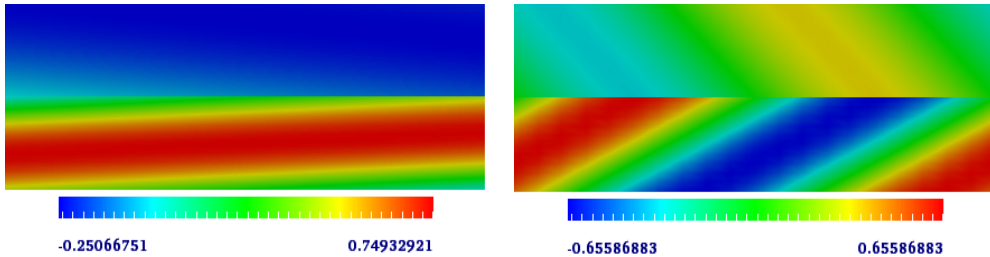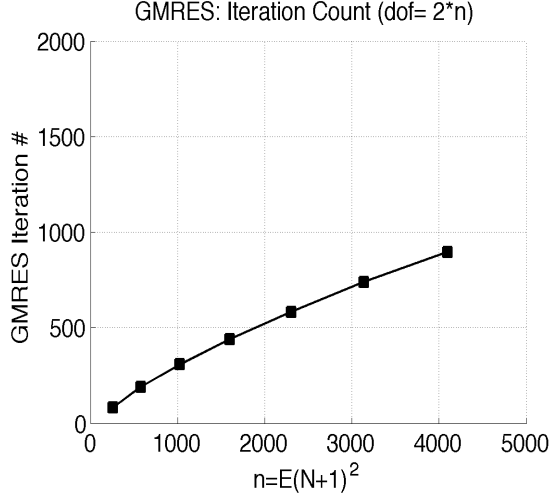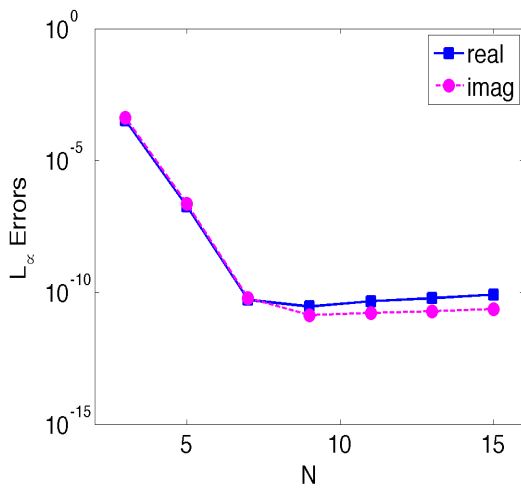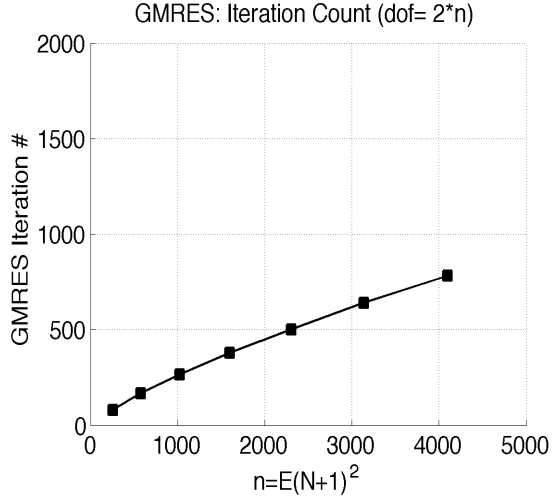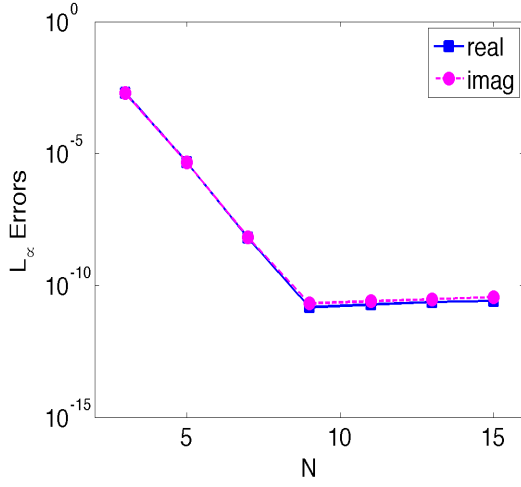
Figure 4: Single layer: $k = 1.5$ (yellow); DtN (top) and Dirichlet (bottom) boundaries.

(a) Mesh and GLL grids with $E = 4 \times 4$, $N = 8$

(b) Real part of the scattered field $\mathbf{U}_{\text{scat}}^N$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

(c) Imaginary part of the scattered field $\mathbf{U}_{\text{scat}}^N$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

Figure 5: Double layer: $k = 1.5$ (yellow) and $k = 2.5$ (blue); DtN (top/bottom) boundaries.

(a) Single layer with smooth flat scattering surface

(b) Double layer with smooth flat scattering surface

Figure 6: Convergence, GMRES iteration counts, and mesh with $E=4\times4$ and $N=3,5,7,9,11,13,15$. The approximation order for the Fourier expansion in the DtN operator is $P=5$.

on Γ. Figure 5(a) shows our mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$. The incident field and analytic solution are given as follows:

- On $\Omega^+ = [0, 2\pi] \times [0, 1]$ with $k^+ = 2.5$ and $\beta = \sqrt{(k^+)^2 - \alpha^2}$:

$$
\begin{aligned}
U_{\text{inc}}^{\text{exact}}(x, y) &= e^{\mathbf{i}(\alpha x - \beta y)}, & (4.3) \\
U^{\text{exact}}(x, y) &= e^{\mathbf{i}(\alpha x - \beta y)} + c^+ e^{\mathbf{i}(\alpha x + \beta y)}. & (4.4)
\end{aligned}
$$

- On $\Omega^- = [0, 2\pi] \times [-1, 0]$ with $k^- = 1.5$ and $\beta = \sqrt{(k^-)^2 - \alpha^2}$:

$$
U^{\text{exact}}(x, y) = c^- e^{\mathbf{i}(\alpha x - \beta y)}. \tag{4.5}
$$

Here the constants are defined by $c^- = 2\beta^+/(\beta^+ + \beta^-)$ and $c^+ = (\beta^+ - \beta^-)/(\beta^+ + \beta^-)$ for $\beta^- = \sqrt{(k^-)^2 - \alpha^2}$ and $\beta^+ = \sqrt{(k^+)^2 - \alpha^2}$. We consider incoming incident waves on Ω^+ for $\alpha = 0.1$ and $\alpha = 1.0$ and compute the total field solutions. We note that the total field is continuous across the surface interface Γ_g, so the continuity condition is applied. In Figures 5(b)–5(c), we demonstrate the scattered field solutions $\mathbf{U}_{\text{scat}}^N$.

Convergence Figure 6 demonstrates convergence studies, measured in the maximum error, for the scattered fields in the single- and double-layer media:

$$
error = \|U_{\text{scat}}^{\text{exact}} - \mathbf{U}_{\text{scat}}^N\|_\infty, \tag{4.6}
$$

where $U_{\text{scat}}^{\text{exact}} = U^{\text{exact}} - U_{\text{inc}}^{\text{exact}}$ is the exact solution for the scattered field. The errors show spectral convergence as N increases. The approximation order for the Fourier data used in the DtN operator is $P = 5$. Table 1 shows the condition numbers increasing as N increases, which explains the errors dropping only up to the 1e-10 level at best when using the GMRES algorithm. Figures 6(a)–6(b) demonstrate the iteration count increasing up to ~ 900 for $N = 15$.

Computation In practice, we transform $\mathbf{U}_{\text{inc}}^N$ into $\mathbf{u}_{\text{inc}}^N = e^{-\mathbf{i}\alpha x}\mathbf{U}_{\text{inc}}^N$ and compute the solution of Eq. (2.14) $\mathbf{u}^N$ with periodic boundary treatment in x and other boundary conditions in y. Then, we transform back to $\mathbf{U}^N$ through the relation $\mathbf{U}^N = e^{\mathbf{i}\alpha x}\mathbf{u}^N$. This approach makes our algorithm much simpler, by eliminating additional boundary treatment in the x-direction. The same idea is applied for solving all other example problems presented in the remaining sections.

4.2 Smooth Curved Scattering Surfaces

In this section, we examine single- and double-layer media with smooth curved surface periodic gratings in the x-direction. Dirichlet and DtN boundary conditions are applied in the y-direction. For these configurations, no analytic solutions are available. We validate our results in comparison with the results by the TFE method [10].

Single Layer Consider a computational domain $\Omega = [0, 2\pi] \times [g(x), 1]$ with the scattering surface defined by $\Gamma_g = \{(x, y) \in \Omega | y = g(x)\}$ and a transparent boundary defined on $\Gamma = \{(x, y) \in \Omega | y = 1\}$. We choose a sinusoidal groove $g(x) = \epsilon \cos(x)$ with the grating depth varying with ϵ. We apply the homogeneous Dirichlet boundary condition on Γ_g and the DtN boundary operator on Γ. Figure 7(a) demonstrates the mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$, representing $g(x)$ with surface fitted elements for the case of $\epsilon = 0.1$. We consider the incident field $U_{\text{inc}}(x, y) =$
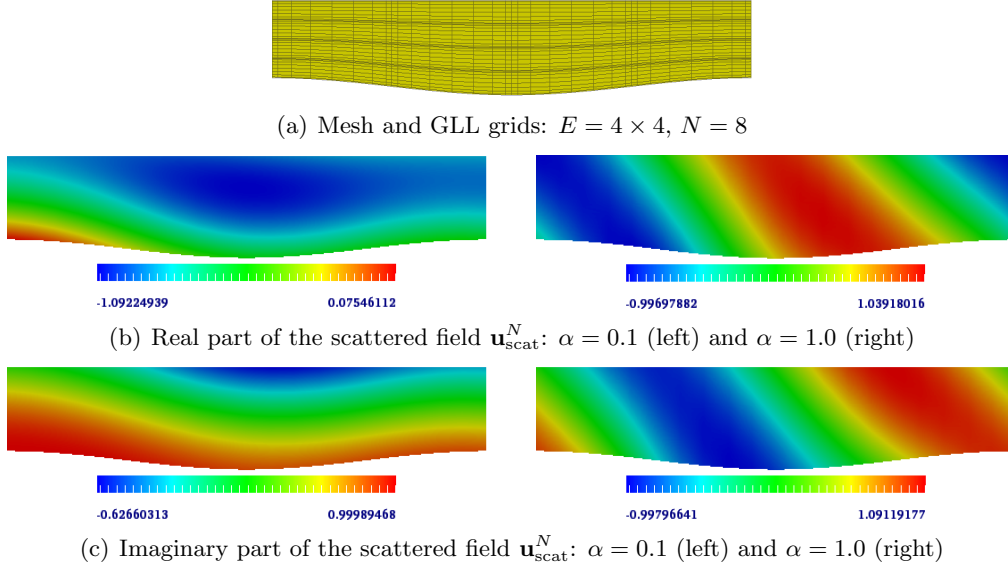
(a) Mesh and GLL grids: $E = 4 \times 4$, $N = 8$

| -1.09224939 | 0.07546112 |

| -0.99697882 | 1.03918016 |

(b) Real part of the scattered field $\mathbf{u}_{\text{scat}}^N$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

| -0.62660313 | 0.99989468 |

| -0.99796641 | 1.09119177 |

(c) Imaginary part of the scattered field $\mathbf{u}_{\text{scat}}^N$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

Figure 7: Single layer: $k = 1.5$ (yellow); DtN (top) and Dirichlet (bottom) boundaries.

(a) Mesh and GLL grids: $E = 4 \times 4$, $N = 8$

| -0.76638919 | 0.74716729 |

| -0.71516794 | 0.70303833 |

(b) Real part of the scattered field $\mathbf{U}_{\text{scat}}^N$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)

| -0.28207076 | 0.75675029 |

| -0.67566937 | 0.71175027 |

(c) Imaginary part of the scattered field $\mathbf{U}_{\text{scat}}^N$: $\alpha = 0.1$ (left) and $\alpha = 1.0$ (right)
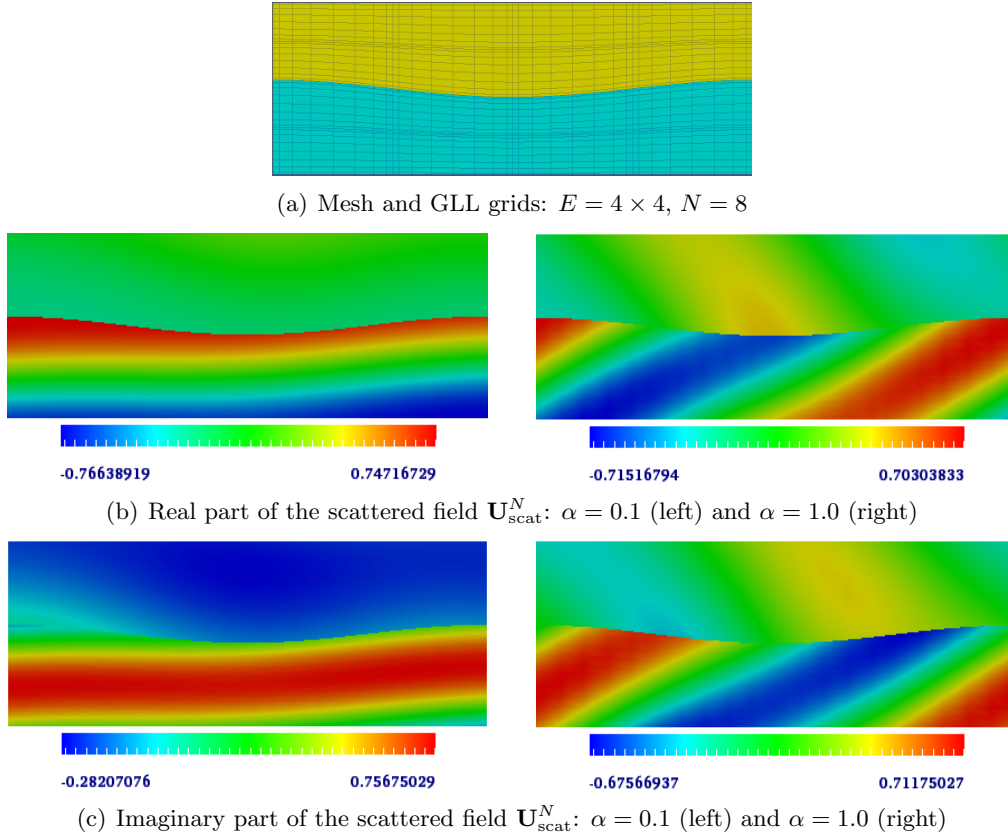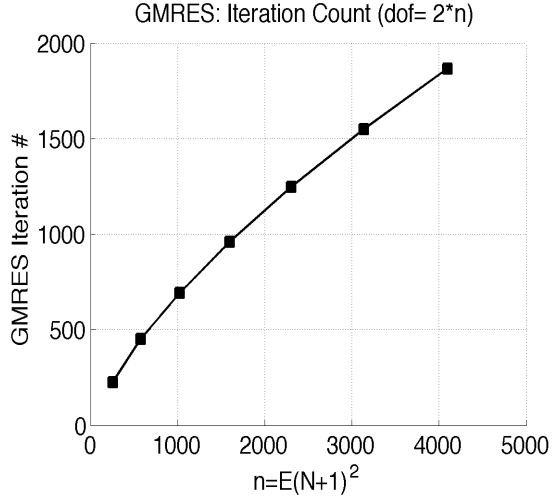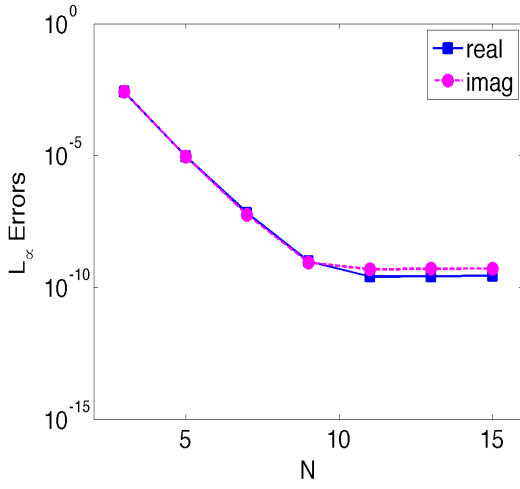
Figure 8: Double layer: $k = 1.5$ (yellow) and $k = 2.5$ (blue); DtN (top/bottom) boundaries.
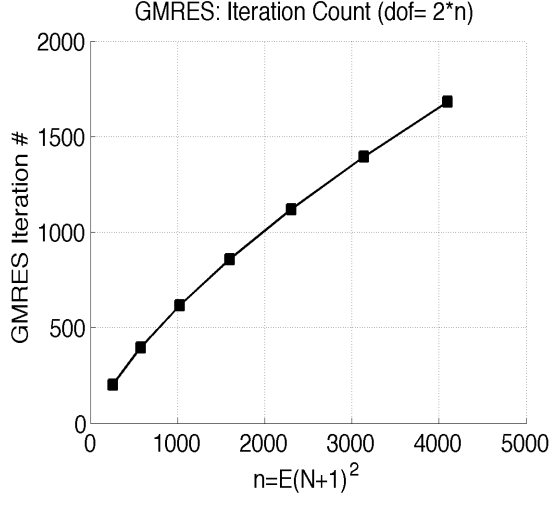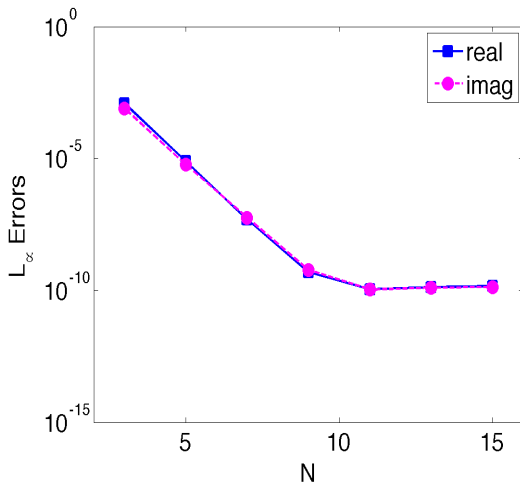
(a) Single layer with smooth curved scattering surface

(b) Double layer with smooth curved scattering surface

Figure 9: Convergence, GMRES iteration counts, and mesh with E=4×4 and N=3,5,7,9,11,13,15. The approximation order for the Fourier expansion in the DtN operator is $P = 5$.

$e^{\mathbf{i}(\alpha x - \beta y)}$ with varying incident angles $\alpha = 0.1$ and $\alpha = 1.0$ for a fixed wavenumber $k = 1.5$, where $\beta = \sqrt{k^2 - \alpha^2}$. We demonstrate the scattered field solutions in Figures 7(b)–7(c).

Double Layer Consider a computational domain $\Omega = \Omega^- \cup \Omega^+$, consisting of two different media $\Omega^+ = [0, 2\pi] \times [g(x), 1]$ and $\Omega^- = [0, 2\pi] \times [-1, g(x)]$ with a sinusoidal groove $g(x) = \epsilon \cos(x)$ at the interface. We define the DtN boundaries on $\Gamma = \Gamma^+ \cup \Gamma^-$ for $\Gamma^+ = \{(x, y) \in \Omega | y = 1\}$ and $\Gamma^- = \{(x, y) \in \Omega | y = -1\}$. Figure 8(a) demonstrates the mesh with $E = 4 \times 4$ and the GLL grids for $N = 8$, representing $g(x)$ with surface-fitted elements for the case of $\epsilon = 0.1$. We consider incoming incident waves $U_{\mathrm{inc}}(x, y) = e^{\mathbf{i}(\alpha x - \beta y)}$ on Ω^+ with varying incident angles of $\alpha = 0.1$ and $\alpha = 1.0$, where $\beta = \sqrt{(k^+)^2 - \alpha^2}$ with $k^+ = 1.5$. We define the wavenumber $k^- = 2.5$ on Ω^-. In Figures 8(b)–8(c), we show the scattered field solutions.

Convergence Figure 9 demonstrates the convergence of our numerical solutions, measured in the maximum error, for the scattered field in the single and double-layer media in comparison with the results by the TFE method:

$$error = \|U_{\mathrm{scat}}^{\mathrm{TFE}} - \mathbf{U}_{\mathrm{scat}}^N\|_\infty, \tag{4.7}$$

where $U_{\mathrm{scat}}^{\mathrm{TFE}}$ is the scattered field solutions by the TFE method. Our solution $\mathbf{U}_{\mathrm{scat}}^N$ on the GLL grids is interpolated on the TFE grids to compute the difference of the solutions on the same grids. The approximation order for the Fourier data used in the DtN operator is $P = 5$. In Figure 9, the errors show spectral convergence as N increases with the GMRES iteration count increasing up to $1700 \sim 1900$ for $N = 15$, as demonstrated in Figures 9(a)–9(b).

4.3 Nonsmooth Scattering Surfaces

In this section, we define the energy defect measure and examine solution behaviors in double-layer media with rectangular, triangular, and sawtooth grooves for the nonsmooth scattering surfaces. We demonstrate the scattered field solutions and the convergence results using the energy defect.

Energy Defect From the grating theory in optics [17], the field outside the groove region can be represented as a series of propagating and evanescent waves, known as the Rayleigh expansion [25]. In general, the Rayleigh expansion is not valid to describe the field inside the groove region. Although the Rayleigh assumption has a limited validity and application, the method has received the attention from many scientists in the related field. Thus we still consider the energy defect measure based on the Rayleigh analysis to validate our results.

Recalling the unbounded subdomains, Ω_0^- and Ω_0^+, defined in (2.5) and some notations from Eqs. (2.15)–(2.19), we can express our solution $U^\pm$ on $\Omega_0^\pm$ as

$$U^+(x, y) = \sum_{p=-\infty}^{\infty} B_p e^{i\alpha_p x + i\beta_p y}, \quad U^-(x, y) = \sum_{q=-\infty}^{\infty} B_q e^{i\alpha_q x + i\beta_q y}, \tag{4.8}$$

where B_p and B_q are the pth and qth Rayleigh amplitudes. The diffraction efficiency is defined by

$$e_p = \frac{\beta_p}{\beta} |B_p|^2 \quad \text{and} \quad e_q = \frac{\beta_q}{\beta} |B_q|^2, \tag{4.9}$$

which satisfy the principle of conservation of energy [23] for lossless media:

$$\sum_{p \in \mathbb{K}} e_p + \sum_{q \in \mathbb{K}} e_q = 1. \tag{4.10}$$

24

To measure the errors of our numerical solutions, we define the energy defect as the following:

$$\varepsilon_{\text{defect}} = \left| 1 - \left(\sum_{p \in \mathbb{K}} \tilde{e}_p + \sum_{q \in \mathbb{K}} \tilde{e}_q \right) \right|, \tag{4.11}$$

where $\tilde{e}_p$ and $\tilde{e}_q$ are the numerical values of e_p and e_q, respectively.

Double Layer We consider a computational domain $\Omega = \Omega^- \cup \Omega^+$ with $\Omega^+ = [0, 2\pi] \times [g(x), 1]$ and $\Omega^- = [0, 2\pi] \times [-1, g(x)]$, including rectangular, triangular, and sawtooth grooves for the scattering surface $g(x)$, as shown in Figures 10–11. The DtN boundaries are defined on $\Gamma = \Gamma^+ \cup \Gamma^-$ for $\Gamma^+ = \{(x, y) \in \Omega | y = 1\}$ and $\Gamma^- = \{(x, y) \in \Omega | y = -1\}$. We consider incoming incident waves $U^{\text{inc}}(x, y) = e^{\mathbf{i}(\alpha x - \beta y)}$ on Ω^+ for varying incident angles of $\alpha = 0$ and $\alpha = 0.2$, where $\beta = \sqrt{(k^+)^2 - \alpha^2}$ with $k^+ = 1.5$. The wavenumber $k^- = 2.5$ is defined on Ω^-. Figures 10–11 show our scattered field solutions. In Table 2, we demonstrate the convergence of our numerical solutions measured in the energy defect, showing spectral convergence as N increases and the GMRES iteration counts increase up to $700 \sim 1400$ for $N = 9$. The approximation order for the Fourier data used in the DtN operator is $P = 5$.

Table 2: Convergence of the energy defect $\varepsilon_{\text{defect}}$ and GMRES iteration count.

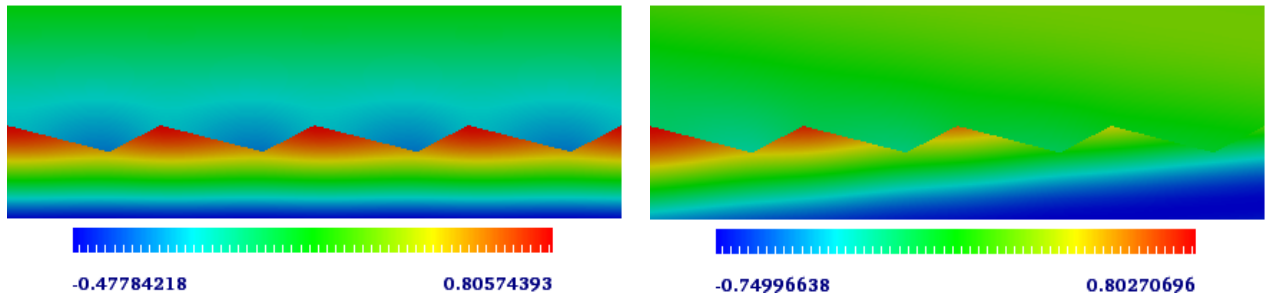| Rectangular Groove							
Normal Incident $\alpha = 0$				Oblique Incident $\alpha = 0.2$			
E	N	$\varepsilon_{\text{defect}}$	iter #	E	N	$\varepsilon_{\text{defect}}$	iter #
64	3	0.435242267233201E-03	226	64	3	0.429733011584389E-03	309
	5	0.713931319988130E-06	447		5	0.701378632226984E-06	638
	7	0.496637989992147E-09	704		7	0.488022389935916E-09	1001
	9	0.793883477475295E-12	998		9	0.139609937066155E-11	1412
Triangular Groove							
Normal Incident $\alpha = 0$				Oblique Incident $\alpha = 0.2$			
E	N	$\varepsilon_{\text{defect}}$	iter #	E	N	$\varepsilon_{\text{defect}}$	iter #
48	3	0.696253818432494E-02	160	48	3	0.679565627532674E-02	177
	5	0.481138116928589E-04	321		5	0.470743511331757E-04	349
	7	0.135466868981619E-06	515		7	0.131819941857974E-06	556
	9	0.208358811685609E-09	728		9	0.187316446415671E-09	782
Sawtooth Groove							
Normal Incident $\alpha = 0$				Oblique Incident $\alpha = 0.2$			
E	N	$\varepsilon_{\text{defect}}$	iter #	E	N	$\varepsilon_{\text{defect}}$	iter #
48	3	0.140770920619001E-01	182	48	3	0.136634700043256E-01	186
	5	0.474567719102777E-04	359		5	0.464225820540073E-04	368
	7	0.133787198504839E-06	563		7	0.130213210270290E-06	574
	9	0.194617803354428E-09	803		9	0.182961906878679E-09	813

5 Conclusions

We studied quasi-periodic solutions of the scalar Helmholtz equation in two dimensions for exterior scattering problems defined on single- and double-layer media with periodic surface grating struc-
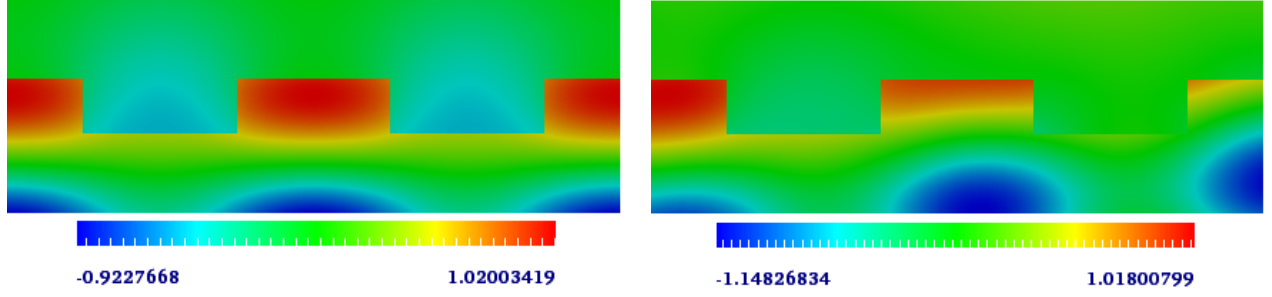
(a) Rectangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (64, 7)$
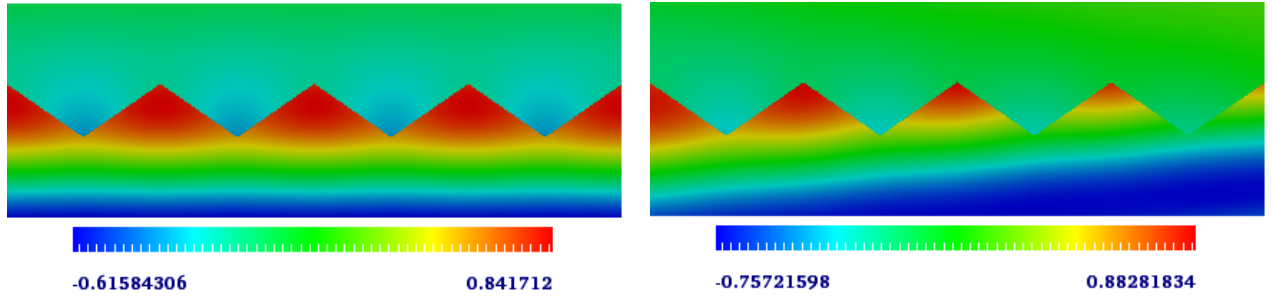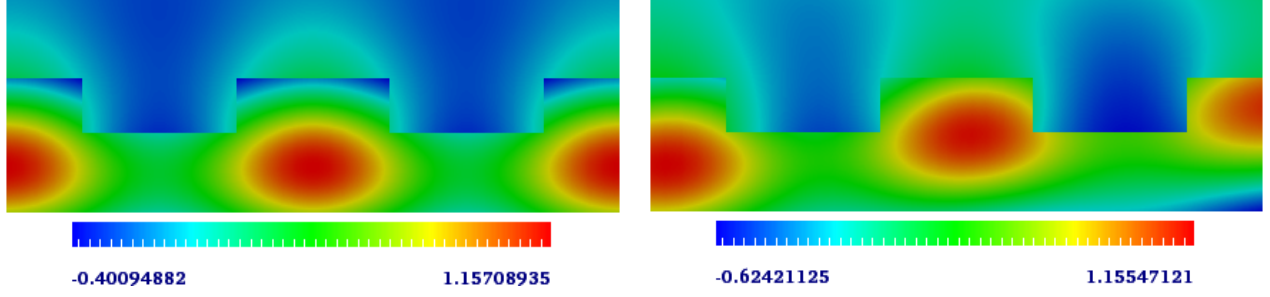

(b) Triangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E,N)=(48,7)$
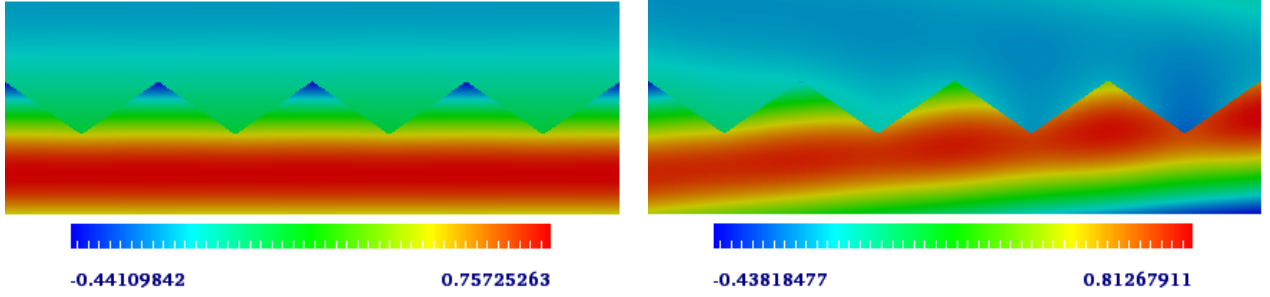

(c) Sawtooth groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (48, 7)$
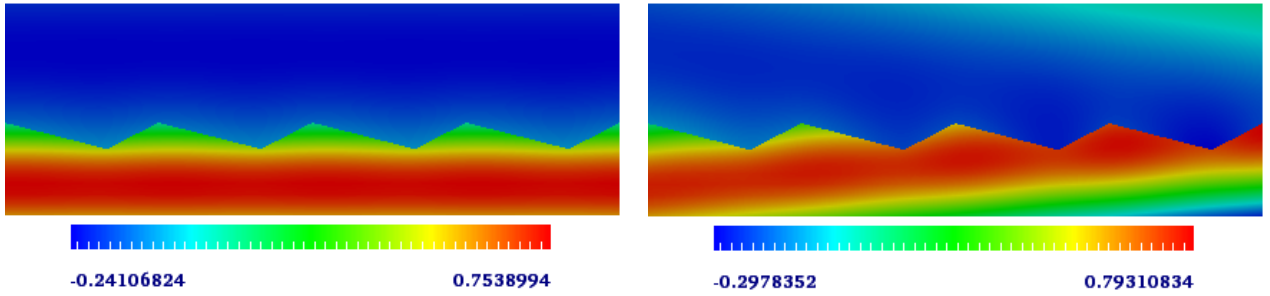
Figure 10: Real part of the scattered field solutions.

(a) Rectangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (64, 7)$

(b) Triangular groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (48, 7)$

(c) Sawtooth groove: $\alpha = 0.0$ (left) and $\alpha = 0.2$ (right); $(E, N) = (48, 7)$

Figure 11: Imaginary part of the scattered field solutions.

tures. We used body-fitted quadrilateral element meshes with spectral element discretization based on the GLL grids. For nonreflecting outgoing waves on the truncated computational domain boundaries, we introduced an accurate formulation of the spectral element DtN boundary operator by representing the Fourier data in relation to the Bessel function—instead of computing the Fourier coefficients using the GLL quadrature integration, which can cause loss of accuracy depending on the grid resolution. Because of the quasi-periodicity and DtN boundary operator, the resulting linear system does not have any special property such as Hermitian positive definite. Thus we applied the GMRES algorithm for solving the linear system. We demonstrated our computational results for the scattered field profiles and their validation with convergence studies showing spectral convergence. Our future efforts include the algorithm extension to three dimensions and the development of efficient preconditioning techniques with reduced iterations for large-scale simulations.

Acknowledgments

References

[1] Prasanta Kumar Banerjee and Roy Butterfield. *Boundary element methods in engineering science*. McGraw-Hill, UK, 1981.

[2] Alex Barnett and Leslie Greengard. A new integral representation for quasi-periodic fields and its application to two-dimensional band structure calculations. *J. Comp. Phys.*, 229:6898–6914, 2010.

[3] Marc Bonnet. *Boundary integral equation methods for solids and fluids*. Wiley, 1999.

[4] Oscar P. Bruno and Fernando Reitich. Numerical solution of diffraction problems: A method of variation of boundaries. *J. Opt. Soc. Am. A*, 10(6):1168–1175, 1993.

[5] Oscar P. Bruno and Fernando Reitich. Numerical solution of diffraction problems: A method of variation of boundaries. II. Finitely conducting gratings, Padé approximants, and singularities. *J. Opt. Soc. Am. A*, 10(11):2307–2316, 1993.

[6] Oscar P. Bruno and Fernando Reitich. Numerical solution of diffraction problems: A method of variation of boundaries. III. Doubly periodic gratings. *J. Opt. Soc. Am. A*, 10(12):2551–2562, 1993.

[7] M.O. Deville, P.F. Fischer, and E.H. Mund. *High-order methods for incompressible fluid flow*. Cambridge University Press, Cambridge, 500 p., 2002.

[8] Qirong Fang, David P. Nicholls, and Jie Shen. A stable, high–order method for three–dimensional bounded–obstacle scattering. *J. Comput. Phys.*, 224(2):1145–1169, 2007.

[9] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73(2):325–348, 1987.

[10] Ying He, David P. Nicholls, and Jie Shen. An efficient and stable spectral method for electromagnetic scattering from a layered periodic structure. *J. Comput. Phys.*, 231(8):3007–3022, 2012.

[11] D. Michael Milder. An improved formalism for rough-surface scattering of acoustic and electromagnetic waves. In *Proceedings of SPIE - The International Society for Optical Engineering (San Diego, 1991)*, volume 1558, pages 213–221. Int. Soc. for Optical Engineering, Bellingham, WA, 1991.

[12] D. Michael Milder. An improved formalism for wave scattering from rough surfaces. *J. Acoust. Soc. Am.*, 89(2):529–541, 1991.

[13] D. Michael Milder. An improved formalism for electromagnetic scattering from a perfectly conducting rough surface. *Radio Science*, 31(6):1369–1376, 1996.

[14] D. Michael Milder. Role of the admittance operator in rough-surface scattering. *J. Acoust. Soc. Am.*, 100(2):759–768, 1996.

[15] D. Michael Milder and H. Thomas Sharp. Efficient computation of rough surface scattering. In *Mathematical and numerical aspects of wave propagation phenomena (Strasbourg, 1991)*, pages 314–322. SIAM, Philadelphia, PA, 1991.

[16] D. Michael Milder and H. Thomas Sharp. An improved formalism for rough surface scattering. ii: Numerical trials in three dimensions. *J. Acoust. Soc. Am.*, 91(5):2620–2626, 1992.

[17] Michel Neviere and Evgeny Popov. *Light propagation in periodic media: Diferential theory and design.* Marcel Dekker Inc., New York, 2003.

[18] David P. Nicholls and Fernando Reitich. A new approach to analyticity of Dirichlet-Neumann operators. *Proc. Roy. Soc. Edinburgh Sect. A*, 131(6):1411–1433, 2001.

[19] David P. Nicholls and Fernando Reitich. Stability of high-order perturbative methods for the computation of Dirichlet-Neumann operators. *J. Comput. Phys.*, 170(1):276–298, 2001.

[20] David P. Nicholls and Fernando Reitich. Analytic continuation of Dirichlet-Neumann operators. *Numer. Math.*, 94(1):107–146, 2003.

[21] David P. Nicholls and Jie Shen. A stable, high–order method for two–dimensional bounded–obstacle scattering. *SIAM J. Sci. Comput.*, 28(4):1398–1419, 2006.

[22] David P. Nicholls and Jie Shen. A rigorous numerical analysis of the transformed field expansion method. *SIAM Journal on Numerical Analysis*, 47(4):2708–2734, 2009.

[23] Roger Petit, editor. *Electromagnetic theory of gratings.* Springer-Verlag, Berlin, 1980.

[24] Youcef Saad and Martin H. Schultz. A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.*, 7(3):856–869, 1986.

[25] J. W. Strutt and Lord Rayleigh. *On the manufacture and theory of diffraction gratings*, volume 74. 1874.

The following paragraph should be deleted before the paper is published: The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory ("Argonne"). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.